

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:27:21 ; Search time 19.4444 Seconds
(without alignment)
15.232 Million cell updates/sec

Title: US-10-057-890A-8

Perfect score: 7

Sequence: 1 TGEKPYK 7

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 90058

Minimum DB seq length: 0
Maximum DB seq length: 10

Post-processing: Listing first 150 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/prodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/prodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/prodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/prodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6	85.7	10	1	US-08-040-548-26
2	6	85.7	10	1	US-08-466-344-26
3	5	71.4	5	3	US-09-173-941-111
4	5	71.4	5	3	US-09-253-396A-231
5	5	71.4	5	3	US-08-676-318A-32
6	5	71.4	5	4	US-09-240-179-49
7	5	71.4	5	4	US-09-229-007A-2
8	5	71.4	5	4	US-09-260-629-13
9	5	71.4	5	4	US-09-731-558-3
10	5	71.4	5	4	US-09-229-037-5
11	5	71.4	8	1	US-08-040-548-20
12	5	71.4	8	1	US-08-466-344-20
13	5	71.4	9	4	US-09-260-629-16
14	5	71.4	10	1	US-08-040-548-24
15	5	71.4	10	1	US-08-466-344-24
16	5	71.4	10	2	US-08-563-148E-2
17	4	57.1	4	4	US-09-260-629-20
18	4	57.1	7	2	US-08-510-215A-4
19	4	57.1	7	2	US-08-510-215A-12
20	4	57.1	9	2	US-08-718-661-7
21	3	42.9	4	1	US-08-456-424-27
22	3	42.9	4	1	US-08-456-424-28
23	3	42.9	4	3	US-08-338-579A-33
24	3	42.9	4	3	US-08-338-579A-92
25	3	42.9	4	3	US-08-863-813A-32
26	3	42.9	4	5	PCT-US94-09851-33
27	3	42.9	4	5	PCT-US94-09851-91

28	3	42.9	4	5	PCT-US94-09851-92
29	3	42.9	5	1	US-08-201-046A-2
30	3	42.9	5	2	US-08-476-062A-51
31	3	42.9	5	3	US-08-208-264A-58
32	3	42.9	5	3	US-08-338-579A-91
33	3	42.9	5	4	US-09-421-097-58
34	3	42.9	5	4	US-09-341-391A-5
35	3	42.9	5	4	US-09-240-179-57
36	3	42.9	5	5	PCT-US96-01314-51
37	3	42.9	5	5	PCT-US96-01314-78
38	3	42.9	5	6	5284931-13
39	3	42.9	6	1	US-08-159-617-2
40	3	42.9	6	1	US-08-288-728-22
41	3	42.9	6	1	US-08-014-979-85
42	3	42.9	6	1	US-08-363-311-2
43	3	42.9	6	1	US-08-406-347A-21
44	3	42.9	6	2	US-08-463-288A-2
45	3	42.9	6	2	US-08-472-244-12
46	3	42.9	6	2	US-08-672-805-18
47	3	42.9	6	2	US-08-470-445A-2
48	3	42.9	6	2	US-08-776-613-3
49	3	42.9	6	2	US-08-776-613-7
50	3	42.9	6	2	US-08-776-613-11
51	3	42.9	6	2	US-08-462-679-2
52	3	42.9	6	2	US-08-466-210A-2
53	3	42.9	6	2	US-08-676-378-8
54	3	42.9	6	2	US-08-467-147A-2
55	3	42.9	6	2	US-08-737-239A-6
56	3	42.9	6	2	US-08-469-014-2
57	3	42.9	6	2	US-08-819-444-3
58	3	42.9	6	2	US-08-819-444-4
59	3	42.9	6	2	US-08-819-444-5
60	3	42.9	6	2	US-08-819-444-6
61	3	42.9	6	2	US-08-819-444-9
62	3	42.9	6	2	US-08-819-444-13
63	3	42.9	6	2	US-08-819-444-14
64	3	42.9	6	2	US-08-937-271-23
65	3	42.9	6	3	US-08-669-408B-11
66	3	42.9	6	3	US-09-334-254A-6
67	3	42.9	6	4	US-09-455-679-42
68	3	42.9	6	4	US-09-508-542-18
69	3	42.9	6	4	US-09-346-230-2
70	3	42.9	6	4	US-08-914-479A-17
71	3	42.9	6	4	US-09-316-919-1
72	3	42.9	6	4	US-09-508-524-18
73	3	42.9	6	5	PCT-US93-10506A-2
74	3	42.9	6	5	PCT-US93-10506-2
75	3	42.9	6	5	PCT-US93-12679-11
76	3	42.9	6	5	PCT-US93-02131-8
77	3	42.9	7	1	US-08-036-555B-5
78	3	42.9	7	1	US-08-469-569-5
79	3	42.9	7	1	US-08-249-322A-5
80	3	42.9	7	1	US-08-040-548-41
81	3	42.9	7	1	US-08-555-669-18
82	3	42.9	7	1	US-08-466-344-41
83	3	42.9	7	1	US-07-923-72A-29
84	3	42.9	7	1	US-07-923-72A-38
85	3	42.9	7	1	US-08-568-147B-12
86	3	42.9	7	1	US-08-469-526A-5
87	3	42.9	7	2	US-08-609-426A-29
88	3	42.9	7	2	US-08-609-426A-38
89	3	42.9	7	2	US-08-374-652C-21
90	3	42.9	7	2	US-08-467-603-30
91	3	42.9	7	2	US-08-734-591A-5
92	3	42.9	7	2	US-08-469-660-5
93	3	42.9	7	2	US-08-466-793-30
94	3	42.9	7	2	US-08-491-861A-30
95	3	42.9	7	2	US-08-310-912A-82
96	3	42.9	7	3	US-09-073-663-18
97	3	42.9	7	3	US-08-841-089-82
98	3	42.9	7	3	US-08-470-335-5
99	3	42.9	7	3	US-08-735-021-5
100	3	42.9	7	3	US-08-734-664A-5

Sequence 92, Appl
Sequence 2, Appl
Sequence 51, Appl
Sequence 58, Appl
Sequence 91, Appl
Sequence 58, Appl
Sequence 5, Appl
Sequence 57, Appl
Sequence 51, Appl
Sequence 78, Appl
Patent No. 5284931
Sequence 2, Appl
Sequence 22, Appl
Sequence 85, Appl
Sequence 21, Appl
Sequence 2, Appl
Sequence 12, Appl
Sequence 18, Appl
Sequence 2, Appl
Sequence 3, Appl
Sequence 7, Appl
Sequence 11, Appl
Sequence 2, Appl
Sequence 8, Appl
Sequence 2, Appl
Sequence 6, Appl
Sequence 2, Appl
Sequence 3, Appl
Sequence 5, Appl
Sequence 5, Appl
Sequence 9, Appl
Sequence 13, Appl
Sequence 14, Appl
Sequence 23, Appl
Sequence 11, Appl
Sequence 6, Appl
Sequence 42, Appl
Sequence 18, Appl
Sequence 2, Appl
Sequence 17, Appl
Sequence 1, Appl
Sequence 18, Appl
Sequence 2, Appl
Sequence 11, Appl
Sequence 8, Appl
Sequence 5, Appl
Sequence 5, Appl
Sequence 41, Appl
Sequence 18, Appl
Sequence 41, Appl
Sequence 29, Appl
Sequence 38, Appl
Sequence 21, Appl
Sequence 30, Appl
Sequence 5, Appl
Sequence 30, Appl
Sequence 30, Appl
Sequence 82, Appl
Sequence 18, Appl
Sequence 82, Appl
Sequence 5, Appl
Sequence 5, Appl
Sequence 5, Appl

Sequence 5, Appli
Sequence 82, Appl
Sequence 80, Appl
Sequence 10, Appl
Sequence 802, Appl
Sequence 889, Appl
Sequence 961, Appl
Sequence 1034, Ap
Sequence 1241, Ap
Sequence 1355, Ap
Sequence 1510, Ap
Sequence 1525, Ap
Sequence 1640, Ap
Sequence 1788, Ap
Sequence 1859, Ap
Sequence 1936, Ap
Sequence 2061, Ap
Sequence 2144, Ap
Sequence 2215, Ap
Sequence 2291, Ap
Sequence 2530, Ap
Sequence 2545, Ap
Sequence 2560, Ap
Sequence 2563, Ap
Sequence 2575, Ap
Sequence 2691, Ap
Sequence 2892, Ap
Sequence 2922, Ap
Sequence 3248, Ap
Sequence 3263, Ap
Sequence 3278, Ap
Sequence 3497, Ap
Sequence 3512, Ap
Sequence 3590, Ap
Sequence 212, App
Sequence 299, App
Sequence 5, Appli
Sequence 30, Appl
Sequence 802, App
Sequence 889, App
Sequence 961, App
Sequence 1034, Ap
Sequence 1241, Ap
Sequence 1355, Ap
Sequence 1510, Ap
Sequence 1525, Ap
Sequence 1640, Ap
Sequence 1788, Ap
Sequence 1859, Ap
Sequence 1936, Ap

ALIGNMENTS

US-08-040-548-26
Sequence 26, Application US/08040548
Patent No. 5763209
GENERAL INFORMATION:
APPLICANT: Sukhatme, Vikas P.
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO THE
FUNCTIONAL DOMAINS OF DNA BINDING PROTEINS
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: 321 No. 5763209th Clark Street, Suite 800
CITY: Chicago
STATE: Illinois
COUNTRY: U.S.A.
ZIP: 60610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/040,548
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Coughlin, Daniel F.
REGISTRATION NUMBER: 36,111
REFERENCE/DOCKET NUMBER: arcd067
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 744-0090
TELEFAX: (312) 245-4961
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-040-548-26

Query Match 85.7%; Score 6; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKPY 6
Db 1 TGEKPY 6

RESULT 2
US-08-466-344-26
Sequence 26, Application US/08466344
Patent No. 5773583
GENERAL INFORMATION:
APPLICANT: Sukhatme, Vikas P.
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO THE
FUNCTIONAL DOMAINS OF DNA BINDING PROTEINS
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: 321 No. 5773583th Clark Street, Suite 800
CITY: Chicago
STATE: Illinois
COUNTRY: U.S.A.
ZIP: 60610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,344
FILING DATE: 06-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/040,548
FILING DATE: 31-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Coughlin, Daniel F.
REGISTRATION NUMBER: 36,111
REFERENCE/DOCKET NUMBER: arcd067
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 744-0090
TELEFAX: (312) 245-4961
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide
US-08-466-344-26

Query Match 85.7%; Score 5; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.34; Indels 0;
Matches 6; Conservative 0; Mismatches 0; Gaps 0;

Qy 1 TGEKP 6
| | | | |
Db 1 TGEKP 6

RESULT 3

US-09-173-941-111
; Sequence 111, Application US/09173941
; Patent No. 6140081
; GENERAL INFORMATION:
; APPLICANT: BARBAS, Carlos F.
; TITLE OF INVENTION: ZINC FINGER BINDING DOMAINS FOR GNN
; FILE REFERENCE: NOV0815
; CURRENT APPLICATION NUMBER: US/09/173.941
; CURRENT FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 111
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: linker
US-09-173-941-111

Query Match 71.4%; Score 5; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05; Indels 0;
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

Qy 1 TGEKP 5
| | | | |
Db 1 TGEKP 5

RESULT 4

US-09-253-396A-231
; Sequence 231, Application US/09253396A
; Patent No. 6205404
; GENERAL INFORMATION:
; APPLICANT: Genome Dynamics, Inc.
; TITLE OF INVENTION: DNA-Binding Proteins of the Zinc-Finger Class
; FILE REFERENCE: 1116242-0003 file: genome03f.app
; CURRENT APPLICATION NUMBER: US/09/253.396A
; CURRENT FILING DATE: 1999-02-19
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 231
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-253-396A-231

Query Match 71.4%; Score 5; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05; Indels 0;
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

Qy 1 TGEKP 5
| | | | |
Db 1 TGEKP 5

RESULT 5

US-08-676-318A-32
; Sequence 32, Application US/08676318A
; Patent No. 6242568
; GENERAL INFORMATION:

APPLICANT: THE SCRIPPS RESEARCH INSTITUTE
; TITLE OF INVENTION: ZINC FINGER PROTEIN DERIVATIVES AND
; METHODS THEREFOR
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/676.318A
; FILING DATE: 30-DEC-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A., Ph.D.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 08401/002001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-678-5070
; TELEFAX: 619-678-5099
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..5
US-08-676-318A-32

Query Match 71.4%; Score 5; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05; Indels 0;
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

Qy 1 TGEKP 5
| | | | |
Db 1 TGEKP 5

RESULT 6

US-09-240-179-49
; Sequence 49, Application US/09240179
; Patent No. 6410248
; GENERAL INFORMATION:
; APPLICANT: Greisman, Harvey A.
; APPLICANT: Pabo, Carl O.
; APPLICANT: Massachusetts Institute of Technology
; TITLE OF INVENTION: A General Strategy for Selecting High-Affinity Zinc
; FINGER PROTEINS FOR DIVERSE DNA TARGET SITES
; FILE REFERENCE: 019496-000220US
; CURRENT APPLICATION NUMBER: US/09/240.179
; CURRENT FILING DATE: 1999-01-29
; EARLIER APPLICATION NUMBER: US 60/073.223
; EARLIER FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 49
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: linker
US-09-240-179-49

Query Match 71.4%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5
|||||
Db 1 TGEKP 5

RESULT 7

US-09-229-007A-2
Sequence 2, Application US/09229007A
Patent No. 6453242
GENERAL INFORMATION:
APPLICANT: Eisenberg, Stephen P.
APPLICANT: Case, Casey C.
APPLICANT: Cox III, George N.
APPLICANT: Jamieson, Andrew
APPLICANT: Rebar, Edward J.
APPLICANT: Sangamo Biosciences, Inc.
TITLE OF INVENTION: Selection of Sites for Targeting by Zinc Finger
TITLE OF INVENTION: Proteins and Methods of Designing Zinc Finger Proteins
TITLE OF INVENTION: co Bind to Preslected Sites
FILE REFERENCE: 019496-001800US
CURRENT APPLICATION NUMBER: US/09/229,007A
CURRENT FILING DATE: 1999-01-12
NUMBER OF SEQ ID NOS: 97
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:peptide linker

US-09-229-007A-2

Query Match 71.4%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5
|||||
Db 1 TGEKP 5

RESULT 8

US-09-260-629-13
Sequence 13, Application US/092606029
Patent No. 6479626
GENERAL INFORMATION:
APPLICANT: Kim, Jin-Soo
APPLICANT: Pabo, Carl O.
APPLICANT: Massachusetts Institute of Technology
TITLE OF INVENTION: Poly Zinc Finger Proteins With Improved Linkers
FILE REFERENCE: 019496-002510US
CURRENT APPLICATION NUMBER: US/09/260,629
CURRENT FILING DATE: 1999-03-01
EARLIER APPLICATION NUMBER: US 60/076,454
EARLIER FILING DATE: 1998-03-02
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 13
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:canonical
OTHER INFORMATION: "TGEKP" linker

US-09-260-629-13

Query Match 71.4%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 71.4%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5
|||||
Db 1 TGEKP 5

RESULT 9

US-09-731-558-3
Sequence 3, Application US/09731558
Patent No. 6503717
GENERAL INFORMATION:
APPLICANT: Case, Casey Christopher
APPLICANT: Liu, Qiang
APPLICANT: Rebar, Edward J.
APPLICANT: Sangamo Biosciences, Inc.
TITLE OF INVENTION: Methods of Using Randomized Libraries of Zinc Finger
TITLE OF INVENTION: Proteins for the Identification of Gene Function
FILE REFERENCE: 019496-003210US
CURRENT APPLICATION NUMBER: US/09/731,558
CURRENT FILING DATE: 2000-12-06
PRIOR APPLICATION NUMBER: US 09/456,100
PRIOR FILING DATE: 1999-12-06
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:flexible
OTHER INFORMATION: polypeptide linker

US-09-731-558-3

Query Match 71.4%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5
|||||
Db 1 TGEKP 5

RESULT 10

US-09-229-037-5
Sequence 5, Application US/09229037A
Patent No. 6534261
GENERAL INFORMATION:
APPLICANT: Cox III, George No. 6534261bert
APPLICANT: Case, Casey Christopher
APPLICANT: Eisenberg, Stephen P.
APPLICANT: Jarvis, Eric Edward
APPLICANT: Spratt, Sharon Kaye
APPLICANT: Sangamo Biosciences, Inc.
TITLE OF INVENTION: Regulation of Endogenous Gene Expression in Cells Using
TITLE OF INVENTION: Zinc Finger Proteins
FILE REFERENCE: 019496-002200US
CURRENT APPLICATION NUMBER: US/09/229,037A
CURRENT FILING DATE: 1999-01-12
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:linker

US-09-229-037-5

Query Match 71.4%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5

Db 1 TGEKP 5
|||||
RESULT 11
US-08-040-548-20
; Sequence 20, Application US/08040548
; Patent No. 5763209
; GENERAL INFORMATION:
; APPLICANT: Sukhatme, Vikas P.
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO THE
; TITLE OF INVENTION: FUNCTIONAL DOMAINS OF DNA BINDING PROTEINS
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: 321 No. 5773583th Clark Street, Suite 800
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60610
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/040,548
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Coughlin, Daniel F.
; REGISTRATION NUMBER: 36,111
; REFERENCE/DOCKET NUMBER: arcd067
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 744-0090
; TELEFAX: (312) 245-4961
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-040-548-20
Query Match 71.4%; Score 5; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGEKP 5
Db 1 TGEKP 5
|||||
RESULT 12
US-08-466-344-20
; Sequence 20, Application US/08466344
; Patent No. 5773583
; GENERAL INFORMATION:
; APPLICANT: Sukhatme, Vikas P.
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO THE
; TITLE OF INVENTION: FUNCTIONAL DOMAINS OF DNA BINDING PROTEINS
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: 321 No. 5773583th Clark Street, Suite 800
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60610
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/040,548
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Coughlin, Daniel F.
; REGISTRATION NUMBER: 36,111
; REFERENCE/DOCKET NUMBER: arcd067
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 744-0090
; TELEFAX: (312) 245-4961
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-040-548-20
Query Match 71.4%; Score 5; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGEKP 5
Db 1 TGEKP 5
|||||

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,344
FILING DATE: 06-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/040,548
FILING DATE: 31-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Coughlin, Daniel F.
REGISTRATION NUMBER: 36,111
REFERENCE/DOCKET NUMBER: arcd067
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 744-0090
TELEFAX: (312) 245-4961
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-466-344-20
Query Match 71.4%; Score 5; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGEKP 5
Db 1 TGEKP 5
|||||
RESULT 13
US-09-260-629-16
; Sequence 16, Application US/09260629
; Patent No. 6479626
; GENERAL INFORMATION:
; APPLICANT: Pabo, Carl O.
; APPLICANT: Massachusetts Institute of Technology
; TITLE OF INVENTION: Poly Zinc Finger Proteins With Improved Linkers
; FILE REFERENCE: 019496-002510US
; CURRENT APPLICATION NUMBER: US/09/260,629
; CURRENT FILING DATE: 1999-03-01
; EARLIER APPLICATION NUMBER: US 60/076,454
; EARLIER FILING DATE: 1998-03-02
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:linker region
US-09-260-629-16
Query Match 71.4%; Score 5; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGEKP 5
Db 2 TGEKP 6
|||||
RESULT 14
US-08-040-548-24
; Sequence 24, Application US/08040548
; Patent No. 5763209
; GENERAL INFORMATION:

APPLICANT: Sukhatme, Vikas P.
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO THE
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: 321 No. 5763209th Clark Street, Suite 800
CITY: Chicago
STATE: Illinois
COUNTRY: U.S.A.
ZIP: 60610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/040,548
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Coughlin, Daniel F.
REGISTRATION NUMBER: 36,111
REFERENCE/DOCKET NUMBER: atcd067
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 744-0090
TELEFAX: (312) 744-4961
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-040-548-24

Query Match 71.4%; Score 5; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGEKP 5
Db 1 TGEKP 5

RESULT 15
US-08-466-344-24
Sequence 24, Application US/08466344
Patent No. 5773583
GENERAL INFORMATION:
APPLICANT: Sukhatme, Vikas P.
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO THE
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: 321 No. 5773583th Clark Street, Suite 800
CITY: Chicago
STATE: Illinois
COUNTRY: U.S.A.
ZIP: 60610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/466,344
FILING DATE: 06-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION NUMBER: 08/040,548
FILING DATE: 31-MAR-1993

ATTORNEY/AGENT INFORMATION:
NAME: Coughlin, Daniel F.
REGISTRATION NUMBER: 36,111
REFERENCE/DOCKET NUMBER: atcd067
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 744-0090
TELEFAX: (312) 245-4961
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-466-344-24

Query Match 71.4%; Score 5; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGEKP 5
Db 1 TGEKP 5

RESULT 16
US-08-563-148E-2
Sequence 2, Application US/08563148E
Patent No. 5929224
GENERAL INFORMATION:
APPLICANT: Kazuo SUZUKI et al.
TITLE OF INVENTION: NOVEL ACTIVATING FACTOR OF LEUKOCYTES
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: 663-2, Shiigi
STREET: Misaki-machi
CITY: Isumi-gun
STATE: Chiba-ken
COUNTRY: JAPAN
ZIP: 299-45
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS v.5
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/563,148E
FILING DATE: No. 5929224ember 27, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 293233/1994
FILING DATE: 28-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: C. Bruce Hamburg
REGISTRATION NUMBER: 22,389
REFERENCE/DOCKET NUMBER: F-5230
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)986-2340
TELEFAX: (212)953-7733
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single strand
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: intermediate fragment
ORIGINAL SOURCE:
ORGANISM: human
TISSUE TYPE: T-cell derived leukemia cells
PUBLICATION INFORMATION:
AUTHORS: Kazuo SUZUKI et al.
TITLE: NOVEL ACTIVATING FACTOR OF LEUKOCYTES
RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 to 10

US-08-563-148E-2

Query Match 71.4%; Score 5; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 EKPYK 7
|
|
|
|
Db 5 EKPYK 9

RESULT 17

US-09-260-629-20
; Sequence 20, Application US/09260629
; Patent No. 6479626
; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Pabo, Carl O.
; APPLICANT: Massachusetts Institute of Technology
; TITLE OF INVENTION: Poly Zinc Finger Proteins With Improved Linkers
; FILE REFERENCE: 019496-002510US
; CURRENT APPLICATION NUMBER: US/09/260,629
; CURRENT FILING DATE: 1999-03-01
; EARLIER APPLICATION NUMBER: US 60/076,454
; EARLIER FILING DATE: 1998-03-02
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:synthetic
; OTHER INFORMATION: linker sequence
US-09-260-629-20

Query Match 57.1%; Score 4; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 GEKP 5
|
|
|
|
Db 1 GEKP 4

RESULT 18

US-08-510-215A-4
; Sequence 4, Application US/08510215A
; Patent No. 5814506
; GENERAL INFORMATION:
; APPLICANT: KONG, Huimin
; APPLICANT: PELLETIER, John J.
; APPLICANT: ALIOTTA, Jason M.
; TITLE OF INVENTION: OVER-EXPRESSION AND PURIFICATION OF
; TITLE OF INVENTION: A TRUNCATED THERMOSTABLE DNA POLYMERASE BY PROTEIN
; TITLE OF INVENTION: FUSION
; CORRESPONDENCE ADDRESS: 14
; NUMBER OF SEQUENCES: 14
; ADDRESS: GREGORY D. WILLIAMS; NEW ENGLAND
; ADDRESSEE: BIO LABS, INC.
; STREET: 32 TOZER ROAD
; CITY: BEVERLY
; STATE: MASSACHUSETTS
; COUNTRY: US
; ZIP: 01915
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/510,215A
; FILING DATE: 02-AUG-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAMS, GREGORY D.
; REGISTRATION NUMBER: 30901
; REFERENCE/DOCKET NUMBER: NEB-113
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 927-5054
; TELEFAX: (508) 927-1705
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-510-215A-4

Query Match 57.1%; Score 4; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAMS, GREGORY D.
; REGISTRATION NUMBER: 30901
; REFERENCE/DOCKET NUMBER: NEB-113
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 927-5054
; TELEFAX: (508) 927-1705
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-510-215A-4

Query Match 57.1%; Score 4; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 GEKP 5
|
|
|
|
Db 3 GEKP 6

RESULT 19

US-08-510-215A-12
; Sequence 12, Application US/08510215A
; Patent No. 5814506
; GENERAL INFORMATION:
; APPLICANT: KONG, Huimin
; APPLICANT: PELLETIER, John J.
; APPLICANT: ALIOTTA, Jason M.
; TITLE OF INVENTION: OVER-EXPRESSION AND PURIFICATION OF
; TITLE OF INVENTION: A TRUNCATED THERMOSTABLE DNA POLYMERASE BY PROTEIN
; TITLE OF INVENTION: FUSION
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS: 14
; ADDRESS: GREGORY D. WILLIAMS; NEW ENGLAND
; ADDRESSEE: BIO LABS, INC.
; STREET: 32 TOZER ROAD
; CITY: BEVERLY
; STATE: MASSACHUSETTS
; COUNTRY: US
; ZIP: 01915
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/510,215A
; FILING DATE: 02-AUG-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAMS, GREGORY D.
; REGISTRATION NUMBER: 30901
; REFERENCE/DOCKET NUMBER: NEB-113
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 927-5054
; TELEFAX: (508) 927-1705
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-510-215A-12

Query Match 57.1%; Score 4; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 GEPK 5
|||||
Db 3 GEPK 6

RESULT 20
US-08-718-661-7
Sequence 7, Application US/08718661
Patent No. 5876972
GENERAL INFORMATION:
APPLICANT: SATTERTHWAIT JR., ARNOLD C.
TITLE OF INVENTION: Nucleic acid molecules coding for mammalian
TITLE OF INVENTION: tumor suppressor proteins and methods for their isolation
NUMBER OF SEQUENCES: 15
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/718.661
FILING DATE:
CLASSIFICATION: 530
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
HYPOTHETICAL: YES
FEATURE:
NAME/KEY: Cross-links
LOCATION: 2
OTHER INFORMATION: /note= "X at position 2 is S or T."

NAME/KEY: Cross-links
LOCATION: 7
OTHER INFORMATION: /note= "X at position 7 is F or Y."

US-08-718-661-7

Query Match 57.1%; Score 4; DB 2; Length 9;
Best Local Similarity 100.0%; Pred.No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 GEPK 5
|||||
Db 3 GEPK 6

RESULT 21
US-08-456-424-27
Sequence 27, Application US/08456424
Patent No. 5807979
GENERAL INFORMATION:
APPLICANT: SATTERTHWAIT JR., ARNOLD C.
APPLICANT: ARRHENIUS, THOMAS
APPLICANT: CABEZAS, EDELMIRA
TITLE OF INVENTION: SYNTHETIC, STABILIZED, THREE-DIMENSION
NUMBER OF SEQUENCES: 145
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 PAGE MILL ROAD
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/456,424
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/224,059
FILING DATE:
APPLICATION NUMBER: US 07/866,040
FILING DATE: 08-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: BOZICEVIC, KARL
REGISTRATION NUMBER: 28,807
REFERENCE/DOCKET NUMBER: 278022000120
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-813-5600
TELEFAX: 415-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-456-424-27

Query Match 42.9%; Score 3; DB 1; Length 4;
Best Local Similarity 100.0%; Pred.No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 TGE 3
|||
Db 2 TGE 4

RESULT 22
US-08-456-424-28
Sequence 28, Application US/08456424
Patent No. 5807979
GENERAL INFORMATION:
APPLICANT: SATTERTHWAIT JR., ARNOLD C.
APPLICANT: ARRHENIUS, THOMAS
APPLICANT: CABEZAS, EDELMIRA
TITLE OF INVENTION: SYNTHETIC, STABILIZED, THREE-DIMENSION
NUMBER OF SEQUENCES: 145
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 PAGE MILL ROAD
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/456,424
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/224,059
FILING DATE:
APPLICATION NUMBER: US 07/866,040
FILING DATE: 08-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: BOZICEVIC, KARL
REGISTRATION NUMBER: 28,807
REFERENCE/DOCKET NUMBER: 278022000120
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-813-5600

TELEFAX: 415-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-456-424-28

Query Match 42.9%; Score 3; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
|||
Db 1 TGE 3

RESULT 23
US-08-338-579A-33

Sequence 33, Application US/08338579A
Patent No. 6068975
GENERAL INFORMATION:
APPLICANT: Gilliam, T. Conrad
APPLICANT: Tanzi, Rudolph E.
TITLE OF INVENTION: ISOLATION AND USES OF A WILSON'S
TITLE OF INVENTION: DISEASE GENE
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/338,579A
FILING DATE: June 17, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/44011-A-PCT-US
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
TELEX:

INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-338-579A-33

Query Match 42.9%; Score 3; DB 3; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
|||
Db 1 TGE 3

RESULT 24

US-08-338-579A-92
Sequence 92, Application US/08338579A
Patent No. 6068975
GENERAL INFORMATION:
APPLICANT: Gilliam, T. Conrad
APPLICANT: Tanzi, Rudolph E.
TITLE OF INVENTION: ISOLATION AND USES OF A WILSON'S
TITLE OF INVENTION: DISEASE GENE
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/338,579A
FILING DATE: June 17, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/44011-A-PCT-US
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
TELEX:

INFORMATION FOR SEQ ID NO: 92:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-338-579A-92

Query Match 42.9%; Score 3; DB 3; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
|||
Db 1 TGE 3

RESULT 25

US-08-863-813A-32
Sequence 32, Application US/08863813A
Patent No. 6140466
GENERAL INFORMATION:
APPLICANT: Barbas III, Carlos F.
APPLICANT: Gottesfeld, Joel M.
APPLICANT: Wright, Peter E.
TITLE OF INVENTION: ZINC FINGER PROTEIN DERIVATIVES
TITLE OF INVENTION: AND METHODS THEREFOR
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/863,813A
FILING DATE: 27-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/676,318
FILING DATE: 18-JUL-1996
APPLICATION NUMBER: 08/183,119
FILING DATE: 18-JAN-1996
APPLICATION NUMBER: US95/00829
FILING DATE: 18-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A., Ph.D.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 08401/010001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5039
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
LOCATION: 4...4
OTHER INFORMATION: where Xaa at position 4 is Lys or Pro

JS-08-863-813A-32

Query Match 42.9%; Score 3; DB 3; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGE 3

Db 1 TGE 3

RESULT 26

PCT-US94-09851-33
Sequence 33, Application PC/TUS9409851
GENERAL INFORMATION:
APPLICANT: Gilliam, T. Conrad
APPLICANT: Tanzi, Rudolph E.
TITLE OF INVENTION: ISOLATION AND USES OF A WILSON'S
TITLE OF INVENTION: DISEASE GENE
NUMBER OF SEQUENCES: 92
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/09851
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/44011-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525

TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
PCT-US94-09851-33

Query Match 42.9%; Score 3; DB 5; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGE 3

Db 1 TGE 3

RESULT 27

PCT-US94-09851-91
Sequence 91, Application PC/TUS9409851
GENERAL INFORMATION:
APPLICANT: Gilliam, T. Conrad
APPLICANT: Tanzi, Rudolph E.
TITLE OF INVENTION: ISOLATION AND USES OF A WILSON'S
TITLE OF INVENTION: DISEASE GENE
NUMBER OF SEQUENCES: 92
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/09851
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/44011-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
PCT-US94-09851-91

Query Match 42.9%; Score 3; DB 5; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGE 3

Db 1 TGE 3

RESULT 28
 PCT-US94-09851-92
 ; Sequence 92, Application PC/TUS9409851
 ; GENERAL INFORMATION:
 ; APPLICANT: Gilliam, T. Conrad
 ; APPLICANT: Tanzi, Rudolph E.
 ; TITLE OF INVENTION: ISOLATION AND USES OF A WILSON'S
 ; TITLE OF INVENTION: DISEASE GENE
 ; NUMBER OF SEQUENCES: 92
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Cooper & Dunham
 ; STREET: 30 Rockefeller Plaza
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: United States of America
 ; ZIP: 10112
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US94/09851
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: White, John P.
 ; REGISTRATION NUMBER: 28,678
 ; REFERENCE/DOCKET NUMBER: 0575/44011-PCT
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 977-9550
 ; TELEFAX: (212) 664-0525
 ; TELEX: 422523 COOP UI
 ; INFORMATION FOR SEQ ID NO: 92:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; PCT-US94-09851-92

Query Match 42.9%; Score 3; DB 5; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TGE 3
 Db 1 TGE 3
 RESULT 29
 US-08-201-046A-2
 ; Sequence 2, Application US/08201046A
 ; Patent No. 5545719
 ; GENERAL INFORMATION:
 ; APPLICANT: Shashoua, Victor E.
 ; TITLE OF INVENTION: NERVE GROWTH PEPTIDES AND USES THEREFOR
 ; NUMBER OF SEQUENCES: 29
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
 ; STREET: 600 Atlantic Ave.
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02210
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/201,046A
 ; FILING DATE: 24-FEB-1994
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Plumer, Elizabeth R.
 ; REGISTRATION NUMBER: 36,637
 ; REFERENCE/DOCKET NUMBER: N0260/7013
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617-720-3500
 ; TELEFAX: 617-720-2441
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 5 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; FRAGMENT TYPE: internal
 ; US-08-201-046A-2
 Query Match 42.9%; Score 3; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 2 GEX 4
 Db 1 GEX 3
 RESULT 30
 US-08-476-062A-51
 ; Sequence 51, Application US/08476062A
 ; Patent No. 5877275
 ; GENERAL INFORMATION:
 ; APPLICANT: Arnaut, M. Amin
 ; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
 ; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS
 ; NUMBER OF SEQUENCES: 53
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson P.C.
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: US
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows95
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/476,062A
 ; FILING DATE: 07-JUN-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/216,081
 ; FILING DATE: 21-MAR-1994
 ; APPLICATION NUMBER: 07/637,830
 ; FILING DATE: 04-JAN-1991
 ; APPLICATION NUMBER: 07/539,842
 ; FILING DATE: 18-JUN-1990
 ; APPLICATION NUMBER: 07/212,573
 ; FILING DATE: 28-JUN-1988
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Freeman, John W.
 ; REGISTRATION NUMBER: 29,066
 ; REFERENCE/DOCKET NUMBER: 00786/068003
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617/542-5070
 ; TELEFAX: 617/542-8906
 ; TELEX: 200154
 ; INFORMATION FOR SEQ ID NO: 51:
 ; SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-476-062A-51

Query Match 42.9%; Score 3; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GEX 4
 |||
 Db 2 GEX 4

RESULT 31

US-08-208-264A-58
 Sequence 58, Application US/08208264A
 Patent No. 6057287

GENERAL INFORMATION:

APPLICANT: MARKLAND, William
 APPLICANT: LADNER, Robert C.
 TITLE OF INVENTION: KALLIKREIN-BINDING "KUNITZ DOMAIN"
 TITLE OF INVENTION: PROTEINS AND ANALOGUES THEREOF
 NUMBER OF SEQUENCES: 61
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Yankwich & Associates
 STREET: 130 Bishop Allen Drive
 CITY: Cambridge
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02139

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5-inch diskette
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: Microsoft Windows 95
 SOFTWARE: Microsoft Word 97
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/208,264A
 FILING DATE: 10-MAR-1994
 CLASSIFICATION: 530

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/179,964
 FILING DATE: 11-JAN-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Leon R. Yankwich
 REGISTRATION NUMBER: 30,237
 REFERENCE/DOCKET NUMBER: DYX-006.1 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-491-4343
 TELEFAX: 617-491-8801

INFORMATION FOR SEQ ID NO: 58:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-208-264A-58

Query Match 42.9%; Score 3; DB 3; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
 |||
 Db 3 TGE 5

RESULT 32

US-08-338-579A-91
 Sequence 91, Application US/08338579A
 Patent No. 6068975

GENERAL INFORMATION:
 APPLICANT: Gilliam, T. Conrad
 APPLICANT: Tanzi, Rudolph E.
 TITLE OF INVENTION: ISOLATION AND USES OF A WILSON'S
 TITLE OF INVENTION: DISEASE GENE
 NUMBER OF SEQUENCES: 107
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Cooper & Dunham
 STREET: 1185 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: United States of America
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/338,579A
 FILING DATE: June 17, 1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: White, John P.
 REGISTRATION NUMBER: 28,678
 REFERENCE/DOCKET NUMBER: 0575/44011-A-PCT-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 278-0400
 TELEFAX: (212) 391-0525
 TELEX:
 INFORMATION FOR SEQ ID NO: 91:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-08-338-579A-91

Query Match 42.9%; Score 3; DB 3; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
 |||
 Db 2 TGE 4

RESULT 33

US-09-421-097-58
 Sequence 58, Application US/09421097
 Patent No. 6333402
 GENERAL INFORMATION:
 APPLICANT: MARKLAND, William
 APPLICANT: LADNER, Robert C.
 TITLE OF INVENTION: KALLIKREIN-BINDING "KUNITZ DOMAIN"
 TITLE OF INVENTION: PROTEINS AND ANALOGUES THEREOF
 NUMBER OF SEQUENCES: 61
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Yankwich & Associates
 STREET: 130 Bishop Allen Drive
 CITY: Cambridge
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02139
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5-inch diskette
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: Microsoft Windows 98
 SOFTWARE: Microsoft Word 97
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/421,097
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION NUMBER: US 08/208,264
FILING DATE: 10-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/179,964
FILING DATE: 11-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Leon R. Yankwich
REGISTRATION NUMBER: 30,237
REFERENCE/DOCKET NUMBER: DYX-006.1 US-1
TELEPHONE: 617-491-4343
TELEFAX: 617-491-8801
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-421-097-58

Query Match 42.9%; Score 3; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGE 3
Db 3 TGE 5

RESULT 34
US-09-341-391A-5
Sequence 5, Application US/09341391A
Patent No. 6342394
GENERAL INFORMATION:
APPLICANT: Rees Smith, Bernard
TITLE OF INVENTION: Diagnosis of Autoimmune Adrenal Disease
FILE REFERENCE: 0769.00117
CURRENT APPLICATION NUMBER: US/09/341,391A
CURRENT FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 9725105.2
PRIOR FILING DATE: 1997-11-28
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 5
TYPE: PRT
ORGANISM: Homo sapiens
US-09-341-391A-5

Query Match 42.9%; Score 3; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 PYK 7
Db 1 PYK 3

RESULT 35
US-09-240-179-57
Sequence 57, Application US/09240179
Patent No. 6410248
GENERAL INFORMATION:
APPLICANT: Greisman, Harvey A.
APPLICANT: Pabo, Carl O.
TITLE OF INVENTION: A General Strategy for Selecting High-Affinity Zinc

TITLE OF INVENTION: Finger Proteins for Diverse DNA Target Sites
FILE REFERENCE: 019496-000220US
CURRENT APPLICATION NUMBER: US/09/240,179
CURRENT FILING DATE: 1999-01-29
EARLIER APPLICATION NUMBER: US 60/073,223
EARLIER FILING DATE: 1998-01-30
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 57
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: linker
US-09-240-179-57

Query Match 42.9%; Score 3; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGE 3
Db 1 TGE 3

RESULT 36
PCT-US96-01314-51
Sequence 51, Application PC/TUS9601314
GENERAL INFORMATION:
APPLICANT: M. Amin Arnaut
TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
NUMBER OF SEQUENCES: 78
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/01314
FILING DATE: 30-JAN-96
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/380,167
FILING DATE: 30-JAN-95
ATTORNEY/AGENT INFORMATION:
NAME: John W. Freeman
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/267001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 5
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
PCT-US96-01314-51

Query Match 42.9%; Score 3; DB 5; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GEX 4
Db 1

Db 2 GEK 4

RESULT 37

PCT-US96-01314-78
Sequence 78, Application PC/TUS9601314
GENERAL INFORMATION:
APPLICANT: M. Amin Arnaout
TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
NUMBER OF SEQUENCES: 78
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/01314
FILING DATE: 30-JAN-96
PRIOR APPLICATION NUMBER: 08/380,167
FILING DATE: 30-JAN-95
ATTORNEY/AGENT INFORMATION:
NAME: John W. Freeman
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/267001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 78:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-01314-78

Query Match 42.9%; Score 3; DB 5; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0;

QY 2 GEK 4

Db 2 GEK 4

RESULT 38

S284931-13
Patent No. 5284931
APPLICANT: SPRINGER, TIMOTHY A.; ROTHLEIN, ROBERT; MARLIN,
STEVEN D.; DUSTIN, MICHAEL L.
TITLE OF INVENTION: INTERCELLULAR ADHESION MOLECULES AND
THEIR BINDING LIGANDS
NUMBER OF SEQUENCES: 41
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/515,478
FILING DATE: 27-APR-1990
SEQ ID NO: 13
LENGTH: 5

S284931-13

Query Match 42.9%; Score 3; DB 6; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0;

QY 2 GEK 4

Db 2 GEK 4

RESULT 39

US-08-159-617-2
Sequence 2, Application US/08159617
Patent No. 5407916
GENERAL INFORMATION:
APPLICANT: Wise, Lawrence D
APPLICANT: Wustrow, David J
TITLE OF INVENTION: Neutotensin Mimetics as
TITLE OF INVENTION: Central Nervous System Agents
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Warner-Lambert Company
STREET: 2800 Plymouth Rd.
CITY: Ann Arbor
STATE: MI
COUNTRY: US
ZIP: 48105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/159,617
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/175556
FILING DATE: 16-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: Tinney, Francis J
REGISTRATION NUMBER: 33069
REFERENCE/DOCKET NUMBER: PD-4411-P1-01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 313 996-7295
TELEFAX: 313 996-1553
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1..2
OTHER INFORMATION: /note= "CH2NH - Bond
OTHER INFORMATION: replaces CONH bond"
US-08-159-617-2

Query Match 42.9%; Score 3; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0;

QY 4 KPY 6

Db 2 KPY 4

RESULT 40

US-08-288-728-22
Sequence 22, Application US/08288728
Patent No. 5493008
GENERAL INFORMATION:
APPLICANT: FOX, JAY W
APPLICANT: TIMPL, RUPERT
TITLE OF INVENTION: TWO NON-CONTIGUOUS REGIONS CONTRIBUTE TO

0; Gaps 0;

```
; TITLE OF INVENTION: NIDOGEN BINDING TO A SINGLE EGF-LIKE MOTIF OF THE LAMININ
; TITLE OF INVENTION: GAMMA-1 CHAIN
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVIS HWY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 15-AUG-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 494-138-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-288-728-22

Query Match 42.9%; Score 3; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
DB 2 TGE 4

RESULT 41
US-08-014-979-85
; Sequence 85, Application US/08014979
; Patent No. 5510240
; GENERAL INFORMATION:
; APPLICANT: Lam, Kit S. et al.
; TITLE OF INVENTION: Random Bio-Oligomer Library, A Method of
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 19930208
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7156-041

; TITLE OF INVENTION: NIDOGEN BINDING TO A SINGLE EGF-LIKE MOTIF OF THE LAMININ
; TITLE OF INVENTION: GAMMA-1 CHAIN
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVIS HWY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 15-AUG-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 494-138-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-288-728-22

Query Match 42.9%; Score 3; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
DB 2 TGE 4

RESULT 42
US-08-363-311-2
; Sequence 2, Application US/08363311
; Patent No. 6548241
; GENERAL INFORMATION:
; APPLICANT: Michel, James L.
; APPLICANT: Ausubel, Frederick M.
; TITLE OF INVENTION: Conjugate Vaccine Against Group B
; TITLE OF INVENTION: Streptococcus
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-2678
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/968,865
; FILING DATE: 02-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 0609.3740004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 466-0800
; TELEFAX: (202) 833-8716
; TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: peptide
; US-08-363-311-2

Query Match 42.9%; Score 3; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
DB 1 GEX 3
```

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 85:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-014-979-85

Query Match 42.9%; Score 3; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GEX 4
DB 1 GEX 3

RESULT 42
US-08-363-311-2
; Sequence 2, Application US/08363311
; Patent No. 6548241
; GENERAL INFORMATION:
; APPLICANT: Michel, James L.
; APPLICANT: Ausubel, Frederick M.
; TITLE OF INVENTION: Conjugate Vaccine Against Group B
; TITLE OF INVENTION: Streptococcus
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-2678
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/968,865
; FILING DATE: 02-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 0609.3740004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 466-0800
; TELEFAX: (202) 833-8716
; TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: peptide
; US-08-363-311-2

Query Match 42.9%; Score 3; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
DB 1 GEX 3
```

Db 4 TGE 6

RESULT 43
US-08-406-347A-21

; Sequence 21, Application US/08406347A

; Patent No. 5714374

; GENERAL INFORMATION:

; APPLICANT: ARNOLD, Edward V.

; APPLICANT: ARNOLD, Gail F.

; TITLE OF INVENTION: CHIMERIC RHINOVIRUSES

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Keryon & Kenyon

; STREET: 1025 Connecticut Avenue, N.W., Suite 600

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.50 inch floppy disk, 1.44 mb storage

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WordPerfect 6.1 Windows

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/406,347A

; FILING DATE: 17-MAR-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/304,635

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/041,790

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/583,335

; ATTORNEY/AGENT INFORMATION:

; NAME: Toffenetti, Judith L.

; REGISTRATION NUMBER: 39,048

; REFERENCE/DOCKET NUMBER: 1984/46203

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-429-1776

; TELEFAX: 202-429-0796

; INFORMATION FOR SEQ ID NO: 21:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 6 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-406-347A-21

Query Match 42.9%; Score 3; DB 1; Length 6;

Best Local Similarity 100.0%; Pred. No. 2.5e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0;

Qy 1 TGE 3

Db 4 TGE 6

RESULT 44

US-08-463-288A-2

; Sequence 2, Application US/08463288A

; Patent No. 5820860

; GENERAL INFORMATION:

; APPLICANT: Michel, James L.

; APPLICANT: Kasper, Dennis L.

; APPLICANT: Ausubel, Frederick M.

; APPLICANT: Madoff, Lawrence C.

; TITLE OF INVENTION: Conjugate Vaccine For Group B

; TITLE OF INVENTION: Streptococcus

; NUMBER OF SEQUENCES: 65

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

; STREET: 1100 New York Avenue, NW, Suite 600

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20005-3934

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/463,288A

; FILING DATE: 05-JUN-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/363,311

; FILING DATE: 22-DEC-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/968,866

; FILING DATE: 02-NOV-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/408,036

; FILING DATE: 15-SEP-1989

; ATTORNEY/AGENT INFORMATION:

; NAME: Bugalsky, Lawrence B.

; REGISTRATION NUMBER: 35,086

; REFERENCE/DOCKET NUMBER: 0609.2370007

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 371-2600

; TELEFAX: (202) 371-2540

; TELEX: 248636 SSK

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 6 amino acids

; TYPE: amino acid

; TOPOLOGY: both

; MOLECULE TYPE: peptide

US-08-463-288A-2

Query Match 42.9%; Score 3; DB 2; Length 6;

Best Local Similarity 100.0%; Pred. No. 2.5e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGE 3

Db 4 TGE 6

RESULT 45

US-08-472-244-12

; Sequence 12, Application US/08472244

; Patent No. 5821088

; GENERAL INFORMATION:

; APPLICANT: Darzins, Aldis

; APPLICANT: Whitehead, Stephen

; APPLICANT: Hruby, Dennis E.

; TITLE OF INVENTION: Use of Gram-Positive Bacteria to Express

; TITLE OF INVENTION: Recombinant Proteins

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, LLP

; STREET: 699 Prince Street

; CITY: Alexandria

; STATE: VA

; COUNTRY: USA

; ZIP: 22314-3187

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/472,244

; FILING DATE: 07-JUN-1995

```
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 016921-065
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..6
; OTHER INFORMATION: /note= "consensus region of SEQ ID NO:1"
US-08-472-244-12

Query Match 42.9%; Score 3; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
Db 4 TGE 6

RESULT 46
US-08-672-805-18
; Sequence 18, Application US/06672805
; Patent No. 5831003
; GENERAL INFORMATION:
; APPLICANT: Baumbach, George A.,
; APPLICANT: Buettner, Joseph A.,
; APPLICANT: Dadd, Christopher A.,
; APPLICANT: Hammond, David J.
; TITLE OF INVENTION: Peptides Which Bind to Prothrombin and
; TITLE OF INVENTION: Thrombin
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bayer Corporation
; STREET: 800 Dwight Way
; STREET: P. O. Box 1986
; CITY: Berkeley
; STATE: California
; COUNTRY: USA
; ZIP: 94701-1986
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb Storage
; COMPUTER: IBM
; OPERATING SYSTEM: DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/06/672,805
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Gibling, James A.
; REGISTRATION NUMBER: 25772
; REFERENCE/DOCKET NUMBER: MSB-7236
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510)705-7910
; TELEFAX: (510)705-7904
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6
```

```
; TYPE: amino acid
; STRANDEDNESS: single strand
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: peptide
US-08-672-805-18

Query Match 42.9%; Score 3; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PYK 7
Db 3 PYK 5

RESULT 47
US-08-470-445A-2
; Sequence 2, Application US/08470445A
; Patent No. 5843444
; GENERAL INFORMATION:
; APPLICANT: Michel, James L.
; APPLICANT: Kasper, Dennis L.
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Madoff, Lawrence C.
; TITLE OF INVENTION: Conjugate Vaccine Against Group B
; TITLE OF INVENTION: Streptococcus
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, NW, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,445A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/363,311
; FILING DATE: 22-DEC-1994
; APPLICATION NUMBER: US 07/969,866
; FILING DATE: 02-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/408,036
; FILING DATE: 15-SEP-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Bugalsky, Lawrence B.
; REGISTRATION NUMBER: 35,086
; REFERENCE/DOCKET NUMBER: 0609,237000A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; TELETYPE: 248636 SSK
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: peptide
US-08-470-445A-2

Query Match 42.9%; Score 3; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


QY 1 TGE 3
DB 4 TGE 6

RESULT 48
US-08-776-613-3
; Sequence 3, Application US/08776613
; Patent No. 5846940
; GENERAL INFORMATION:
; APPLICANT: Okamoto, Sinseiro
; APPLICANT: Amagaya, Sakae
; APPLICANT: Sakamoto, Kenji
; APPLICANT: Kanitani, Masanao
; TITLE OF INVENTION: CORNEAL THERAPEUTIC AGENT
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 No. 5846940th Glebe Rd. 8th floor
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201-4741
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,613
; FILING DATE: 03-FEB-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP95/01549
; FILING DATE: 04-AUG-1995
; NAME: Crawford, Arthur R.
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 160-293
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-776-613-3

Query Match 42.9%; Score 3; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0;

QY 4 KPY 6
DB 2 KPY 4

RESULT 49
US-08-776-613-7
; Sequence 7, Application US/08776613
; Patent No. 5846940
; GENERAL INFORMATION:
; APPLICANT: Okamoto, Sinseiro
; APPLICANT: Amagaya, Sakae
; APPLICANT: Sakamoto, Kenji
; APPLICANT: Kanitani, Masanao
; TITLE OF INVENTION: CORNEAL THERAPEUTIC AGENT

NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 No. 5846940th Glebe Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/776,613
FILING DATE: 03-FEB-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP95/01549
FILING DATE: 04-AUG-1995
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 160-293
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-776-613-7

Query Match 42.9%; Score 3; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0;

QY 4 KPY 6
DB 2 KPY 4

RESULT 50
US-08-776-613-11
; Sequence 11, Application US/08776613
; Patent No. 5846940
; GENERAL INFORMATION:
; APPLICANT: Okamoto, Sinseiro
; APPLICANT: Amagaya, Sakae
; APPLICANT: Sakamoto, Kenji
; APPLICANT: Kanitani, Masanao
; TITLE OF INVENTION: CORNEAL THERAPEUTIC AGENT
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 No. 5846940th Glebe Rd. 8th floor
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201-4741
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,613

; FILING DATE: 03-FEB-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP95/01549
; FILING DATE: 04-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-184844
; FILING DATE: 05-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford, Arthur R.
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 160-293
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-776-613-11

Query Match 42.9%; Score 3; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0;

Qy 4 KPY 6
Db 2 KPY 4

RESULT 51

US-08-462-679-2
; Sequence 2, Application US/08462679
; Patent No. 5847081
; GENERAL INFORMATION:
; APPLICANT: Michel, James L.
; APPLICANT: Kasper, Dennis L.
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Madoff, Lawrence C.
; TITLE OF INVENTION: Conjugate Vaccine For Group B
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, NW, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,679
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/363,311
; FILING DATE: 22-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/968,866
; FILING DATE: 02-NOV-1992
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-776-613-11

; REFERENCE/DOCKET NUMBER: 0609.2370008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: peptide
; US-08-462-679-2

Query Match 42.9%; Score 3; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0;

Qy 1 TGE 3
Db 4 TGE 6

RESULT 52

US-08-466-210A-2
; Sequence 2, Application US/08466210A
; Patent No. 5858362
; GENERAL INFORMATION:
; APPLICANT: Michel, James L.
; APPLICANT: Kasper, Dennis L.
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Madoff, Lawrence C.
; TITLE OF INVENTION: Conjugate Vaccine For Group B
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, NW, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,210A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/363,311
; FILING DATE: 22-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/968,866
; FILING DATE: 02-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/408,036
; FILING DATE: 15-SEP-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Bugaisky, Lawrence B.
; REGISTRATION NUMBER: 35,086
; REFERENCE/DOCKET NUMBER: 0609.237000B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: peptide

US-08-466-210A-2

Query Match 42.9%; Score 3; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
DB 4 TGE 6

RESULT 53

US-08-676-378-8
; Sequence 8, Application US/08676378
; Patent No. 5858727
; GENERAL INFORMATION:
; APPLICANT: Meens, Jochen
; APPLICANT: Klose, Michael
; APPLICANT: Sahm, Hermann
; APPLICANT: Freudl, Roland
; TITLE OF INVENTION: SECRETION OF OUTER MEMBRANE PROTEINS OF
; TITLE OF INVENTION: GRAM NEGATIVE BACTERIA BY MEANS OF GRAM POSITIVE HOST
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSER: The Firm of Karl F. Ross, PC
; STREET: 5676 Riverdale Ave.
; CITY: Bronx
; STATE: New York
; COUNTRY: USA
; ZIP: 10471-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: IBM PC compatible
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/676,378
; FILING DATE: 18-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Jonathan E
; REGISTRATION NUMBER: 26,963
; REFERENCE/DOCKET NUMBER: 20067
; TELEPHONE: (718) 884-6600
; TELEFAX: 718/601-1099
; TELEX: 620428
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-676-378-8

Query Match 42.9%; Score 3; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KPY 6
DB 1 KPY 3

RESULT 54

US-08-467-147A-2
; Sequence 2, Application US/08467147A
; Patent No. 5908629
; GENERAL INFORMATION:
; APPLICANT: Michel, James L.
; APPLICANT: Kasper, Dennis L.
; APPLICANT: Ausubel, Frederick M.

; APPLICANT: Madoff, Lawrence C.
; TITLE OF INVENTION: Conjugate Vaccine For Group B
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, NW, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,147A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/363,311
; FILING DATE: 22-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/968,866
; FILING DATE: 02-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/408,036
; FILING DATE: 15-SEP-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Bugaisky, Lawrence B.
; REGISTRATION NUMBER: 35,086
; REFERENCE/DOCKET NUMBER: 0609.2370009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: peptide

US-08-467-147A-2

Query Match 42.9%; Score 3; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
DB 4 TGE 6

RESULT 55

US-08-737-299A-6
; Sequence 6, Application US/08737299A
; Patent No. 5952464
; GENERAL INFORMATION:
; APPLICANT: MERTENS et al.
; TITLE OF INVENTION: Labelled Peptide Compounds
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jeffrey S. Boone
; STREET: 675 McDonnell Blvd.
; CITY: St. Louis
; STATE: MO
; COUNTRY: US
; ZIP: 63134
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/737,299A
; FILING DATE: 21-JUL-1997
; CLASSIFICATION: 530
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02111
; FILING DATE: 21-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Jeffrey S. Boone
; REGISTRATION NUMBER: 29284
; REFERENCE/DOCKET NUMBER: M0863-WO-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 1 (314) 654-8955
; TELEFAX: 1 (314) 654-3156
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /product= "(2-I-phenyl)acetyl
; OTHER INFORMATION: arginine, wherein I is 123-I, 125-I or 131-I"
;
; US-08-737-299A-6
;
; Query Match 42.9%; Score 3; DB 2; Length 6;
; Best Local Similarity 100.0%; Pred. No. 2.5e+05;
; Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 4 KPY 6
; DB 2 KPY 4
;
; RESULT 56
; US-08-469-014-2
; Sequence 2, Application US/08469014
; Patent No. 5568521
; GENERAL INFORMATION:
; APPLICANT: Michel, James L.
; APPLICANT: Kasper, Dennis L.
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Madoff, Lawrence C.
; TITLE OF INVENTION: Conjugate Vaccine Against Group B
; TITLE OF INVENTION: Streptococcus
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, NW, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,014
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/363,311
; FILING DATE: 22-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/968,866
; FILING DATE: 02-NOV-1992
```

```
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/408,036
; FILING DATE: 15-SEP-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Bugelsky, Lawrence B.
; REGISTRATION NUMBER: 35,086
; REFERENCE/DOCKET NUMBER: 0609.2370006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: peptide
;
; US-08-469-014-2
;
; Query Match 42.9%; Score 3; DB 2; Length 6;
; Best Local Similarity 100.0%; Pred. No. 2.5e+05;
; Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 TGE 3
; DB 4 TGE 6
;
; RESULT 57
; US-08-819-444-3
; Sequence 3, Application US/08819444A
; Patent No. 5968763
; GENERAL INFORMATION:
; APPLICANT: FISCHETTI, Vincent A.
; APPLICANT: PANCHOLI, Vijaykumar
; TITLE OF INVENTION: ENZYME FOR CLEAVAGE OF THE ANCHOR REGION OF SURFACE
; FILE REFERENCE: 016921-105
; CURRENT APPLICATION NUMBER: US/08/819,444A
; CURRENT FILING DATE: 1997-03-17
; EARLIER APPLICATION NUMBER: US 08/319,540
; EARLIER FILING DATE: 1994-10-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
;
; US-08-819-444-3
;
; Query Match 42.9%; Score 3; DB 2; Length 6;
; Best Local Similarity 100.0%; Pred. No. 2.5e+05;
; Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 TGE 3
; DB 4 TGE 6
;
; RESULT 58
; US-08-819-444-4
; Sequence 4, Application US/08819444A
; Patent No. 5968763
; GENERAL INFORMATION:
; APPLICANT: FISCHETTI, Vincent A.
; APPLICANT: PANCHOLI, Vijaykumar
; TITLE OF INVENTION: ENZYME FOR CLEAVAGE OF THE ANCHOR REGION OF SURFACE
; FILE REFERENCE: 016921-105
; CURRENT APPLICATION NUMBER: US/08/819,444A
; CURRENT FILING DATE: 1997-03-17
; EARLIER APPLICATION NUMBER: US 08/319,540
; EARLIER FILING DATE: 1994-10-07
```

```

; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 6
; TYPE: PRT
; ORGANISM: E. faecalis or S. cremoris
US-08-819-444-4

Query Match
Best Local Similarity 42.9%; Score 3; DB 2; Length 6;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGE 3
Db 4 TGE 6

RESULT 59
US-08-819-444-5
; Sequence 5, Application US/08819444A
; Patent No. 5968763
; GENERAL INFORMATION:
; APPLICANT: FISCHETTI, Vincent A.
; TITLE OF INVENTION: ENZYME FOR CLEAVAGE OF THE ANCHOR REGION OF SURFACE
; FILE REFERENCE: 016921-105
; CURRENT APPLICATION NUMBER: US/08/819,444A
; CURRENT FILING DATE: 1997-03-17
; EARLIER APPLICATION NUMBER: US 08/319,540
; EARLIER FILING DATE: 1994-10-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-08-819-444-5

Query Match
Best Local Similarity 42.9%; Score 3; DB 2; Length 6;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGE 3
Db 4 TGE 6

RESULT 60
US-08-819-444-6
; Sequence 6, Application US/08819444A
; Patent No. 5968763
; GENERAL INFORMATION:
; APPLICANT: FISCHETTI, Vincent A.
; TITLE OF INVENTION: ENZYME FOR CLEAVAGE OF THE ANCHOR REGION OF SURFACE
; FILE REFERENCE: 016921-105
; CURRENT APPLICATION NUMBER: US/08/819,444A
; CURRENT FILING DATE: 1997-03-17
; EARLIER APPLICATION NUMBER: US 08/319,540
; EARLIER FILING DATE: 1994-10-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 6
; TYPE: PRT
; ORGANISM: group G streptococcus
US-08-819-444-6

Query Match
Best Local Similarity 42.9%; Score 3; DB 2; Length 6;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGE 3
Db 4 TGE 6

RESULT 61
US-08-819-444-9
; Sequence 9, Application US/08819444A
; Patent No. 5968763
; GENERAL INFORMATION:
; APPLICANT: FISCHETTI, Vincent A.
; TITLE OF INVENTION: ENZYME FOR CLEAVAGE OF THE ANCHOR REGION OF SURFACE
; FILE REFERENCE: 016921-105
; CURRENT APPLICATION NUMBER: US/08/819,444A
; CURRENT FILING DATE: 1997-03-17
; EARLIER APPLICATION NUMBER: US 08/319,540
; EARLIER FILING DATE: 1994-10-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-08-819-444-9

Query Match
Best Local Similarity 42.9%; Score 3; DB 2; Length 6;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGE 3
Db 4 TGE 6

RESULT 62
US-08-819-444-13
; Sequence 13, Application US/08819444A
; Patent No. 5968763
; GENERAL INFORMATION:
; APPLICANT: FISCHETTI, Vincent A.
; TITLE OF INVENTION: ENZYME FOR CLEAVAGE OF THE ANCHOR REGION OF SURFACE
; FILE REFERENCE: 016921-105
; CURRENT APPLICATION NUMBER: US/08/819,444A
; CURRENT FILING DATE: 1997-03-17
; EARLIER APPLICATION NUMBER: US 08/319,540
; EARLIER FILING DATE: 1994-10-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Staphylococcus mutans
US-08-819-444-13

Query Match
Best Local Similarity 42.9%; Score 3; DB 2; Length 6;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGE 3
Db 4 TGE 6

RESULT 63
US-08-819-444-14
; Sequence 14, Application US/08819444A
; Patent No. 5968763
; GENERAL INFORMATION:
```

APPLICANT: FISCHETTI, Vincent A.
APPLICANT: PANCHOLI, Vijaykumar
TITLE OF INVENTION: ENZYME FOR CLEAVAGE OF THE ANCHOR REGION OF SURFACE
TITLE OF INVENTION: PROTEINS FROM GRAM POSITIVE BACTERIA
FILE REFERENCE: 016921-105
CURRENT APPLICATION NUMBER: US/08/819,444A
EARLIER FILING DATE: 1997-03-17
EARLIER APPLICATION NUMBER: US 08/319,540
EARLIER FILING DATE: 1994-10-07
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 6
TYPE: PRT
ORGANISM: S. mutans or S. pyogenes
US-08-819-444-14

Query Match 42.9%; Score 3; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
DB 4 TGE 6

RESULT 64
US-08-937-271-23
Sequence 23, Application US/08937271
Patent No. 6063386
GENERAL INFORMATION:
APPLICANT: Dale, James B.
APPLICANT: Lederer, James W.
TITLE OF INVENTION: RECOMBINANT MULTIVALENT M PROTEIN
TITLE OF INVENTION: VACCINE
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/937,271
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Rosenman, Stephen J.
REGISTRATION NUMBER: 43,058
REFERENCE/DOCKET NUMBER: 481112.405C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-937-271-23

Query Match 42.9%; Score 3; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
DB 4 TGE 6

us-10-057-890a-8.oligo.ra1

Db 4 TGE 6

RESULT 65
US-08-669-408B-11
Sequence 11, Application US/08669408B
Patent No. 6100055
GENERAL INFORMATION:
APPLICANT: GUSS, Bengt
APPLICANT: JONSSON, Hans
APPLICANT: LINDBERG, Martin
APPLICANT: MUELLER, Hans-Peter
APPLICANT: RANTAMAKI, Liisa K.
TITLE OF INVENTION: METHOD AND MEANS FOR PRODUCING
TITLE OF INVENTION: PLASMAPROTEINASE INHIBITOR-BINDING PROTEINS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/669,408B
FILING DATE: 03-JUL-1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/SE94/00826
FILING DATE: 06-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9302855-3
FILING DATE: 06-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 61743/102
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-669-408B-11

Query Match 42.9%; Score 3; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
DB 4 TGE 6

RESULT 66
US-09-334-254A-6
Sequence 6, Application US/09334254A
Patent No. 6194386
GENERAL INFORMATION:
APPLICANT: MERTENS et al.
TITLE OF INVENTION: Labelled Peptide Compounds
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jeffrey S. Boone

STREET: 675 McDonnell Blvd.
CITY: St. Louis
STATE: MO
COUNTRY: US
ZIP: 63134
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/334,254A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/737,299
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Jeffrey S. Boone
REGISTRATION NUMBER: 29284
REFERENCE/DOCKET NUMBER: M0863-WO-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 1 (314) 654-8955
TELEFAX: 1 (314) 654-3156
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /product= "(2-I-phenyl)acetyl
OTHER INFORMATION: arginine, wherein I is 123-I, 125-I or 131-I"
US-09-334-254A-6

Query Match 42.9%; Score 3; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 KPY 6
Db 2 KPY 4
RESULT 67
US-09-455-679-42
Sequence 42, Application US/09455679
Patent No. 6329186
GENERAL INFORMATION:
APPLICANT: Nielsen, Bjarne R.
APPLICANT: Svendsen, Allan
APPLICANT: Bojsen, Kirsten
APPLICANT: Vind, Jesper
APPLICANT: Pedersen, Henrik
TITLE OF INVENTION: Glucosylases With N-Terminal Extensions
FILE REFERENCE: 5691.200-US
CURRENT APPLICATION NUMBER: US/09/455,679
CURRENT FILING DATE: 1999-12-07
EARLIER APPLICATION NUMBER: PA 1998 01616
EARLIER FILING DATE: 1998-12-07
EARLIER APPLICATION NUMBER: PA 1999 00409
EARLIER FILING DATE: 1999-03-24
EARLIER APPLICATION NUMBER: 60/111,674
EARLIER FILING DATE: 1998-12-10
EARLIER APPLICATION NUMBER: 60/126,740
EARLIER FILING DATE: 1999-03-29
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 42

LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Peptide
US-09-455-679-42
Query Match 42.9%; Score 3; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGE 3
Db 4 TGE 6
RESULT 68
US-09-508-542-18
Sequence 18, Application US/09508542
Patent No. 6339174
GENERAL INFORMATION:
APPLICANT: STRAUSS, ANDREAS
APPLICANT: THUMM, GUNTHER
APPLICANT: POHLNER, JOHANNES
APPLICANT: GOTZ, FRIEDRICH
TITLE OF INVENTION: METHOD FOR IDENTIFYING A NUCLEIC ACID
FILE REFERENCE: 10496/P65266US0
CURRENT APPLICATION NUMBER: US/09/508,542
CURRENT FILING DATE: 2000-05-16
PRIOR APPLICATION NUMBER: PCT/EP98/06136
PRIOR FILING DATE: 1998-09-26
PRIOR APPLICATION NUMBER: 97 116 841.4
PRIOR FILING DATE: 1997-09-27
PRIOR APPLICATION NUMBER: 97 118 755.4
PR-CR FILING DATE: 1997-10-29
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 18
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Motif
US-09-508-542-18
Query Match 42.9%; Score 3; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGE 3
Db 4 TGE 6
RESULT 69
US-09-346-290-2
Sequence 2, Application US/09346290
Patent No. 6342223
GENERAL INFORMATION:
APPLICANT: Michel, James L.
APPLICANT: Kasper, Dennis L.
APPLICANT: Ausubel, Frederick M.
APPLICANT: Madoff, Lawrence C.
TITLE OF INVENTION: Conjugate Vaccine Against Group B
TITLE OF INVENTION: Streptococcus
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, NW, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3934

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/346,290
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA: 08/469,014
FILING DATE:
APPLICATION NUMBER: US 07/968,866
FILING DATE: 02-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/408,036
FILING DATE: 15-SEP-1989
ATTORNEY/AGENT INFORMATION:
NAME: Bugalsky, Lawrence B.
REGISTRATION NUMBER: 35,086
REFERENCE/DOCKET NUMBER: 0609.2370006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
TELEX: 248636 SSK
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: both
MOLECULE TYPE: peptide
US-09-346-290-2

Query Match 42.9%; Score 3; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGE 3
Db 4 TGE 6

RESULT 70
US-08-914-479A-17
Sequence 17, Application US/08914479A
Patent No. 6419932
GENERAL INFORMATION:
APPLICANT: Dale, James B.
TITLE OF INVENTION: ANTIGEN OF HYBRID M PROTEIN AND CARRIER
FILE OF INVENTION: FOR GROUP A STREPTOCOCCAL VACCINE
FILE REFERENCE: 48112.404C2
CURRENT APPLICATION NUMBER: US/08/914,479A
PRIOR FILING DATE: 1997-08-19
PRIOR APPLICATION NUMBER: 08/409,270
PRIOR FILING DATE: 1995-03-23
PRIOR APPLICATION NUMBER: 07/945,860
PRIOR FILING DATE: 1992-09-16
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Hexapeptide consensus sequence
US-08-914-479A-17

Query Match 42.9%; Score 3; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGE 3

Db 4 TGE 6
RESULT 71
US-09-316-919-1
Sequence 1, Application US/09316919
Patent No. 6469154
GENERAL INFORMATION:
APPLICANT: Teien, Roger Y.
TITLE OF INVENTION: FLUORESCENT PROTEIN INDICATORS
FILE REFERENCE: 07257/073001
CURRENT APPLICATION NUMBER: US/09/316,919
CURRENT FILING DATE: 1999-05-21
NUMBER OF SEQ ID NOS: 63
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Linker moiety
US-09-316-919-1

Query Match 42.9%; Score 3; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGE 3
Db 3 TGE 5

RESULT 72
US-09-508-524-18
Sequence 18, Application US/09508524
Patent No. 6566062
GENERAL INFORMATION:
APPLICANT: STRAUSS, ANDREAS
APPLICANT: THUMM, GUNTHER
APPLICANT: FOHLNER, JOHANNES
APPLICANT: GOTZ, FRIEDRICH
TITLE OF INVENTION: METHOD FOR IDENTIFYING A NUCLEIC ACID
FILE REFERENCE: 10496/P65266US0
CURRENT APPLICATION NUMBER: US/09/508,524
CURRENT FILING DATE: 2000-05-02
PRIOR APPLICATION NUMBER: PCT/EP98/06136
PRIOR FILING DATE: 1998-09-26
PRIOR APPLICATION NUMBER: 97 116 841.4
PRIOR FILING DATE: 1997-09-27
PRIOR APPLICATION NUMBER: 97 118 755.4
PRIOR FILING DATE: 1997-10-29
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 18
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Motif
US-09-508-524-18

Query Match 42.9%; Score 3; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGE 3
Db 4 TGE 6

RESULT 73


```

PCT-US93-10506A-2
; Sequence 2, Application PC/TUS9310506A
; GENERAL INFORMATION:
; APPLICANT: THE GENERAL HOSPITAL CORPORATION
; ADDRESS: Fruit Street
; APPLICANT: Boston, Massachusetts 02114
; APPLICANT: United States of America
; APPLICANT: 75 Francis Street
; APPLICANT: Boston, Massachusetts 02115
; APPLICANT: United States of America
; TITLE OF INVENTION: Conjugate Vaccine Against Group B
; TITLE OF INVENTION: Streptococcus
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Ave., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/10506A
; FILING DATE: 02-NOV-1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/968,866
; FILING DATE: 02-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 0609.237PC01
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: peptide
; PCT-US93-10506A-2

Query Match 42.9%; Score 3; DB 5; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGE 3
Db 4 TGE 6

RESULT 74
PCT-US93-10506-2
; Sequence 2, Application PC/TUS9310506
; GENERAL INFORMATION:
; APPLICANT: Michel, James L.
; APPLICANT: Kasper, Dennis
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Madoff, Lawrence C.
; TITLE OF INVENTION: Conjugate Vaccine Against Group B
; TITLE OF INVENTION: Streptococcus
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, N.W.; Suite 600
; CITY: Washington
; STATE: D.C.

```

```

; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/10506
; FILING DATE: 02-NOV-1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/968,866
; FILING DATE: 02-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 0609.237PC01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: peptide
; PCT-US93-10506-2

Query Match 42.9%; Score 3; DB 5; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGE 3
Db 4 TGE 6

RESULT 75
PCT-US93-12679-11
; Sequence 11, Application PC/TUS9312679
; GENERAL INFORMATION:
; APPLICANT: Rath, Matthias
; TITLE OF INVENTION: SYNTHETIC OLIGOPEPTIDES ANALOGOUS TO
; TITLE OF INVENTION: PROTEIN SIGNAL SEQUENCES -METHODS OF IDENTIFICATION AND
; TITLE OF INVENTION: METHODS OF USE
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SHELDON & MAK
; STREET: 401 Florence Street
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/12679
; FILING DATE: 30-DEC-1993
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cranfill, Raymond B
; REGISTRATION NUMBER: 32,845
; REFERENCE/DOCKET NUMBER: RATH-10016PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5333
; TELEFAX: 415-322-5499
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:

```

;
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
PCT-US93-12679-11

Query Match 42.9%; Score 3; DB 5; Length 6;
Best Local Similarity 100.0%; Pred.No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EXP 5
|||
DB 3 EXP 5

RESULT 76

PCT-US95-02131-8
; Sequence 8, Application PC/TUS9502131
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Labelled peptide compounds
; NUMBER OF SEQUENCES: 16
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02131
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 94200409.4
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 1
; OTHER INFORMATION: Xaa is (2-*I-phenyl)acetyl-Arg
PCT-US95-02131-8

Query Match 42.9%; Score 3; DB 5; Length 6;
Best Local Similarity 100.0%; Pred.No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KPY 6
|||
DB 2 KPY 4

RESULT 77

US-08-036-555B-5
; Sequence 5, Application US/08036555B
; Patent No. 5530109
; GENERAL INFORMATION:
; APPLICANT: Goodearl, Andrew; Stroobant, Paul;
; APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;
; APPLICANT: Chen, Maio Su; Hiles, Ian
; TITLE OF INVENTION: Glial Mitogenic Factors, Their
; PREPARATION AND USE
; NUMBER OF SEQUENCES: 184
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City

;
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/036,555B
; FILING DATE: 24-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/965,173
; FILING DATE: 23-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/940,389
; FILING DATE: 03-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/907,138
; FILING DATE: 30-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/863,703
; FILING DATE: 03-APRIL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.K. 91 07566.3
; FILING DATE: 10-APRIL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Teal, Christine H.
; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: LUD 5250.4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; FEATURE:
; OTHER INFORMATION: Xaa in position 1 is Lysine or Arginine.
US-08-036-555B-5

Query Match 42.9%; Score 3; DB 1; Length 7;
Best Local Similarity 100.0%; Pred.No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GEK 4
|||
DB 3 GEK 5

RESULT 78

US-08-469-569-5
; Sequence 5, Application US/08469569
; Patent No. 5606032
; GENERAL INFORMATION:
; APPLICANT: Goodearl, Andrew; Stroobant, Paul;
; APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;
; APPLICANT: Chen, Maio Su; Hiles, Ian
; TITLE OF INVENTION: Glial Mitogenic Factors, Their
; PREPARATION AND USE
; NUMBER OF SEQUENCES: 184
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

```

;
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,569
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/036,555
; FILING DATE: 24-MAR-1993
; APPLICATION NUMBER: 07/965,173
; FILING DATE: 23-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/940,389
; FILING DATE: 03-SEP-1992
; APPLICATION NUMBER: 07/907,138
; FILING DATE: 30-JUN-1992
; APPLICATION NUMBER: 07/863,703
; FILING DATE: 03-APRIL-1992
; APPLICATION NUMBER: U.K. 91 07566.3
; FILING DATE: 10-APRIL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsai, Christine H.
; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: LUD 5250.4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; FEATURE:
; OTHER INFORMATION: Xaa in position 1 is Lysine or Arginine.
US-08-469-569-5
Query Match 42.9%; Score 3; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 GEK 4
Db 3 GEK 5
RESULT 79
US-08-249-322A-5
; Sequence 5, Application US/08249322A
; Patent No. 5716930
; GENERAL INFORMATION:
; APPLICANT: Goodearl, Andrew; Stroobant, Paul;
; APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;
; APPLICANT: Chen, Maio Su; Hiles, Ian
; TITLE OF INVENTION: Glial Mitogenic Factors, Their
; TITLE OF INVENTION: Preparation and Use
; NUMBER OF SEQUENCES: 184
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect

```

```

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/249,322A
; FILING DATE: 26-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/036,555
; FILING DATE: 24-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/965,173
; FILING DATE: 23-OCT-1992
; APPLICATION NUMBER: 07/940,389
; FILING DATE: 03-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/907,138
; FILING DATE: 30-JUN-1992
; APPLICATION NUMBER: 07/863,703
; FILING DATE: 03-APRIL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.K. 91 07566.3
; FILING DATE: 10-APRIL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsai, Christine H.
; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: LUD 250.4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; FEATURE:
; OTHER INFORMATION: Xaa in position 1 is Lysine or Arginine.
US-08-249-322A-5
Query Match 42.9%; Score 3; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 GEK 4
Db 3 GEK 5
RESULT 80
US-08-040-548-41
; Sequence 41, Application US/08040548
; Patent No. 5763209
; GENERAL INFORMATION:
; APPLICANT: Sukhatme, Vikas P.
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO THE
; TITLE OF INVENTION: FUNCTIONAL DOMAINS OF DNA BINDING PROTEINS
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: 321 No. 5763209th Clark Street, Suite 800
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60610
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/040,548
; FILING DATE:
; CLASSIFICATION: 514

```

ATTORNEY/AGENT INFORMATION:
NAME: Coughlin, Daniel F.
REGISTRATION NUMBER: 36,111
REFERENCE/DOCKET NUMBER: arc067
TELEPHONE: (312) 744-0090
TELEFAX: (312) 245-4961
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Xaa
LOCATION: 4
IDENTIFICATION METHOD: Xaa = Arg or Lys
FEATURE:
NAME/KEY: Xaa
LOCATION: 6
IDENTIFICATION METHOD: Xaa = Phe or Tyr
FEATURE:
NAME/KEY: Xaa
LOCATION: 7
IDENTIFICATION METHOD: Xaa = Any amino acid
US-08-040-548-41
Query Match 42.9%; Score 3; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;
QY 1 TGE 3
Db 1 TGE 3
RESULT 81
US-08-555-669-18
Sequence 18, Application US/08555669
Patent No. 5773248
GENERAL INFORMATION:
APPLICANT: Brewton, Richard G.
APPLICANT: Mayne, Richard
TITLE OF INVENTION: TYPE IX COLLAGEN AND FRAGMENTS THEREOF
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/555,669
FILING DATE: 13-NOV-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Halluin, Albert P.
REGISTRATION NUMBER: 25,227
REFERENCE/DOCKET NUMBER: 8389-030
TELEPHONE: 415-854-3660
TELEFAX: 415-854-3694
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids

TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-555-669-18
Query Match 42.9%; Score 3; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;
QY 2 GEK 4
Db 2 GEK 4
RESULT 82
US-08-466-344-41
Sequence 41, Application US/08466344
Patent No. 5773583
GENERAL INFORMATION:
APPLICANT: Sukhatme, Vikas P.
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO THE
FUNCTIONAL DOMAINS OF DNA BINDING PROTEINS
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: 321 No. 5773583th Clark Street, Suite 800
CITY: Chicago
STATE: Illinois
COUNTRY: U.S.A.
ZIP: 60610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,344
FILING DATE: 06-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/040,548
FILING DATE: 31-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Coughlin, Daniel F.
REGISTRATION NUMBER: 36,111
REFERENCE/DOCKET NUMBER: arc067
TELEPHONE: (312) 744-0090
TELEFAX: (312) 245-4961
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Xaa
LOCATION: 4
IDENTIFICATION METHOD: Xaa = Arg or Lys
FEATURE:
NAME/KEY: Xaa
LOCATION: 6
IDENTIFICATION METHOD: Xaa = Phe or Tyr
FEATURE:
NAME/KEY: Xaa
LOCATION: 7
IDENTIFICATION METHOD: Xaa = Any amino acid
US-08-466-344-41
Query Match 42.9%; Score 3; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGE 3
|||
Db 1 TGE 3

RESULT 83

US-07-923-724-29
; Sequence 29, Application US/07923724
; Patent No. 5780292
; GENERAL INFORMATION:
; APPLICANT: Nevalainen, Helena K.M.
; APPLICANT: Paloheimo, Marja T.
; APPLICANT: Miettinen-Oinonen, Arja S.K.
; APPLICANT: Torkkeli, Tuula K.
; APPLICANT: Cantrell, Michael
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Rambosek, John A.
; APPLICANT: Turunen, Marja K.
; APPLICANT: Fagerstr m, Richard B.
; TITLE OF INVENTION: Production of Phytase Degrading Enzymes
; TITLE OF INVENTION: in Trichoderma
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/923,724
; FILING DATE: 31-JUL-1992

CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/496,155
; FILING DATE: 19-MAR-1990

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/044,077
; FILING DATE: 29-APR-1987

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 8610600
; FILING DATE: 30-APR-1986

ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 1050.0240004

TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 29:

SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: both

US-07-923-724-29

Query Match 42.9%; Score 3; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GEX 4
|||
Db 5 GEX 7

RESULT 84

US-07-923-724-38
; Sequence 38, Application US/07923724
; Patent No. 5780292
; GENERAL INFORMATION:
; APPLICANT: Nevalainen, Helena K.M.
; APPLICANT: Paloheimo, Marja T.
; APPLICANT: Miettinen-Oinonen, Arja S.K.
; APPLICANT: Torkkeli, Tuula K.
; APPLICANT: Cantrell, Michael
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Rambosek, John A.
; APPLICANT: Turunen, Marja K.
; APPLICANT: Fagerstr m, Richard B.
; TITLE OF INVENTION: Production of Phytase Degrading Enzymes
; TITLE OF INVENTION: in Trichoderma
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/923,724
; FILING DATE: 31-JUL-1992

CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/496,155
; FILING DATE: 19-MAR-1990

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/044,077
; FILING DATE: 29-APR-1987

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 8610600
; FILING DATE: 30-APR-1986

ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 1050.0240004

TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 38:

SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: both

US-07-923-724-38

Query Match 42.9%; Score 3; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GEX 4
|||
Db 5 GEX 7

RESULT 85

US-08-568-147B-12
; Sequence 12, Application US/08568147B
; Patent No. 5783422
; GENERAL INFORMATION:
; APPLICANT: Suminami, Yoshinori
; APPLICANT: Kato, Hiroshi
; APPLICANT: Sekiguchi, Kiyoshi
; APPLICANT: Takeda, Katsumichi

;; TITLE OF INVENTION: DNA FRAGMENT CODING FOR SQUAMOUS CELL
;; TITLE OF INVENTION: CARCINOMA-ASSOCIATED ANTIGEN
;; NUMBER OF SEQUENCES: 17
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Scully, Scott, Murphy & Presser
;; STREET: 400 Garden City Plaza
;; CITY: Garden City
;; STATE: New York
;; COUNTRY: USA
;; ZIP: 11530
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/568,147B
;; FILING DATE:
;; CLASSIFICATION: 536
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 800,952
;; FILING DATE: 02-DEC-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: DiGiglio, Frank S.
;; REGISTRATION NUMBER: 31,346
;; REFERENCE/DOCKET NUMBER: 8425
;; TELEPHONE: 516-742-4343
;; TELEFAX: 516-742-4366
;; TELEX: 230 901 SANS UR
;; INFORMATION FOR SEQ ID NO: 12:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 7 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-08-568-147B-12

Query Match 42.9%; Score 3; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 PYK 7
Db 5 PYK 7

RESULT 86
US-08-469-526A-5
; Sequence 5, Application US/08469526A
; Patent No. 5792849
; GENERAL INFORMATION:
; APPLICANT: Goodearl, Andrew
; APPLICANT: Stroobant, Paul
; APPLICANT: Mirghetti, Luisa
; APPLICANT: Waterfield, Michael
; APPLICANT: Marchionni, Mark
; APPLICANT: Chen, Maio Su
; APPLICANT: Hiles, Ian
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
; TITLE OF INVENTION: PREPARATION AND USE
; NUMBER OF SEQUENCES: 187
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS

;; SOFTWARE: FastSEQ for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/469,526A
;; FILING DATE: 06 June 1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/036,555
;; FILING DATE: 24-MAR-1993
;; APPLICATION NUMBER: 07/965,173
;; FILING DATE: 23-OCT-1992
;; APPLICATION NUMBER: 07/940,389
;; FILING DATE: 03-SEP-1992
;; APPLICATION NUMBER: 07/907,138
;; FILING DATE: 03-JUN-1992
;; APPLICATION NUMBER: 07/863,703
;; FILING DATE: 03-APRIL-1992
;; APPLICATION NUMBER: U.K. 91 07566.3
;; FILING DATE: 10-APR-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Bieker-Brady, Kristina
;; REGISTRATION NUMBER: 39,109
;; REFERENCE/DOCKET NUMBER: 04585/00200A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 617-428-0200
;; TELEFAX: 617-428-7045
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 7
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; FEATURE:
;; OTHER INFORMATION: Xaa in position 1 is Lysine or
;; OTHER INFORMATION: Arginine.
;; US-08-469-526A-5

Query Match 42.9%; Score 3; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GEK 4
Db 3 GEK 5

RESULT 87
US-08-609-426A-29
; Sequence 29, Application US/08609426A
; Patent No. 5830733
; GENERAL INFORMATION:
; APPLICANT: Nevalainen, Helena K.M.
; APPLICANT: Paloheimo, Marja T.
; APPLICANT: Miettinen-Oinonen, Arja S.K.
; APPLICANT: Torkkeli, Tuula K.
; APPLICANT: Cantrell, Michael
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Rambosek, John A.
; APPLICANT: Turunen, Marja K.
; APPLICANT: Fagerstr m, Richard B.
; APPLICANT: Houston, Christine S.
; TITLE OF INVENTION: Production of Phytase Degrading Enzymes
; TITLE OF INVENTION: in Trichoderma
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,426A
; FILING DATE: 01-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/923,724
; FILING DATE: 31-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/496,155
; FILING DATE: 19-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/044,077
; FILING DATE: 29-APR-1987
; APPLICATION NUMBER: UK 8610600
; FILING DATE: 30-APR-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Reed, Grant E.
; REGISTRATION NUMBER: P-41,264
; REFERENCE/DOCKET NUMBER: 1050.0080001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; US-08-609-426A-29

Query Match 42.9%; Score 3; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GEK 4
Db 5 GEK 7

RESULT 88
US-08-609-426A-38
; Sequence 38, Application US/08609426A
; Patent No. 5830733
; GENERAL INFORMATION:
; APPLICANT: Nevalainen, Helena K.M.
; APPLICANT: Paloheimo, Marja T.
; APPLICANT: Torkkeli, Tuula K.
; APPLICANT: Cantrell, Michael
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Rambosek, John A.
; APPLICANT: Turunen, Marja K.
; APPLICANT: Fagerstr m, Richard B.
; APPLICANT: Houston, Christine S.
; TITLE OF INVENTION: Production of Phytase Degrading Enzymes
; TITLE OF INVENTION: in Trichoderma
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,426A

```

```

; FILING DATE: 01-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/923,724
; FILING DATE: 31-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/496,155
; FILING DATE: 19-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/044,077
; FILING DATE: 29-APR-1987
; APPLICATION NUMBER: UK 8610600
; FILING DATE: 30-APR-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Reed, Grant E.
; REGISTRATION NUMBER: P-41,264
; REFERENCE/DOCKET NUMBER: 1050.0080001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; US-08-609-426A-38

Query Match 42.9%; Score 3; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GEK 4
Db 5 GEK 7

RESULT 89
US-08-374-652C-21
; Sequence 21, Application US/08374652C
; Patent No. 5834286
; GENERAL INFORMATION:
; APPLICANT: NEVALAINEN, HELENA K.M.
; APPLICANT: PALOHEIMO, MARJA T.
; APPLICANT: FAGERSTROM, RICHARD B.
; APPLICANT: MIETTINEN-OINONEN, ARJA S.
; APPLICANT: TURUNEN, MARJA K.
; APPLICANT: RAMBOSEK, JOHN A.
; APPLICANT: PIDDINGTON, CHRISTOPHER S.
; APPLICANT: HOUSTON, CHRISTINE S.
; APPLICANT: CANTRELL, MICHAEL A.
; TITLE OF INVENTION: RECOMBINANT CELLS, DNA CONSTRUCTS,
; TITLE OF INVENTION: VECTORS AND METHODS FOR EXPRESSING PHYTASE DEGRADING
; TITLE OF INVENTION: ENZYMES IN DESIRED RATIOS
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/374,652C
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07058

```

1 FILING DATE: 27-JUL-1993
2 CLASSIFICATION: 435
3 PRIOR APPLICATION DATA:
4 APPLICATION NUMBER: US 07/925,401
5 FILING DATE: 31-JUL-1992
6 CLASSIFICATION: 435
7 ATTORNEY/AGENT INFORMATION:
8 NAME: REED, GRANT E.
9 REGISTRATION NUMBER: 41,264
10 REFERENCE/DOCKET NUMBER: 1050.071001
11 TELECOMMUNICATION INFORMATION:
12 TELEPHONE: 202-371-2600
13 TELEFAX: 202-371-2540
14 INFORMATION FOR SEQ ID NO: 21:
15 SEQUENCE CHARACTERISTICS:
16 LENGTH: 7 amino acids
17 TYPE: amino acid
18 STRANDEDNESS: single
19 TOPOLOGY: not relevant
20 MOLECULE TYPE: peptide
21 US-08-374-652C-21

Query Match 42.9%; Score 3; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GEK 4
Db 5 GEK 7

RESULT 90
US-08-467-603-30
Sequence 30, Application US/08467603
Patent No. 5843672
GENERAL INFORMATION:
APPLICANT: Morgenstern, Jay P.
APPLICANT: Kanieczny, Andrzej
APPLICANT: Bizindaukas, Christine B.
APPLICANT: Brauer, Andrew W.
TITLE OF INVENTION: Allergenic Proteins and
TITLE OF INVENTION: Peptides from Dog
TITLE OF INVENTION: Dander and Uses Therefor
NUMBER OF SEQUENCES: 104
CORRESPONDENCE ADDRESS:
ADDRESSER: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII-text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,603
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/156,549
FILING DATE:
APPLICATION NUMBER: 07/999,712
FILING DATE: 31-Dec-92
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-026CP(IPC-048CP)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 30:

1 SEQUENCE CHARACTERISTICS:
2 LENGTH: 7 amino acids
3 TYPE: amino acid
4 TOPOLOGY: linear
5 MOLECULE TYPE: peptide
6 FRAGMENT TYPE: internal
7 US-08-467-603-30

Query Match 42.9%; Score 3; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EXP 5
Db 5 EXP 7

RESULT 91
US-08-734-591A-5
Sequence 5, Application US/08734591A
Patent No. 5854220
GENERAL INFORMATION:
APPLICANT: Goodearl, Andrew
APPLICANT: Stroobant, Paul
APPLICANT: Minghetti, Luisa
APPLICANT: Waterfield, Michael
APPLICANT: Hiles, Ian
APPLICANT: Marchionni, Mark
APPLICANT: Chen, Mario
TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
TITLE OF INVENTION: PREPARATION AND USE
NUMBER OF SEQUENCES: 187
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM Compatible Pentium
OPERATING SYSTEM: Windows95
SOFTWARE: WordPerfect (Version 7.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/734,591A
FILING DATE: 22-OCT-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/470,335
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/036,555
FILING DATE: 03-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/965,173
FILING DATE: 23-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/940,389
FILING DATE: 03-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/907,138
FILING DATE: 30-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/863,703
FILING DATE: 03-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: UK 91 07566.3
FILING DATE: 10-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Bleker-Brady, Kristina
REGISTRATION NUMBER: 39,109
REFERENCE/DOCKET NUMBER: 04585/00200P

TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 428-0200
TELEFAX: (617) 428-7045
TELEX:

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 7

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

FEATURE:

OTHER INFORMATION: Xaa in position 1 is Lysine or Arginine.

US-08-734-591A-5

Query Match 42.9%; Score 3; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GEK 4
Db 3 GEK 5

RESULT 92

US-08-469-660-5

Sequence 5, Application US/08469660

Patent No. 5876973

GENERAL INFORMATION:

APPLICANT: Gwynne, David I.; Marchionni, Mark;

APPLICANT: McBurney, Robert N.

TITLE OF INVENTION: INHIBITORS OF CELL PROLIFERATION,

TITLE OF INVENTION: THEIR PREPARATION AND USE

NUMBER OF SEQUENCES: 184

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson

STREET: 225 Franklin Street

CITY: Boston

STATE: Massachusetts

ZIP: 0211-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/469,660

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/011,396

FILING DATE: 29-JAN-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/984,085

FILING DATE: 01-DEC-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/951,747

FILING DATE: 25-SEP-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/927,337

FILING DATE: 10-AUG-1992

ATTORNEY/AGENT INFORMATION:

NAME: Clark, Paul T.

REGISTRATION NUMBER: 30,162

REFERENCE/DOCKET NUMBER: 04585/017004

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 542-5070

TELEFAX: 200154

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 7

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

FEATURE:
OTHER INFORMATION: Xaa in position 1 is Lysine or
OTHER INFORMATION: Arginine.
US-08-469-660-5

Query Match 42.9%; Score 3; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0;

QY 2 GEK 4
Db 3 GEK 5

RESULT 93

US-08-466-793-30

Sequence 30, Application US/08466793

Patent No. 5891716

GENERAL INFORMATION:

APPLICANT: Morgenstern, Jay P.

APPLICANT: Kanieczny, Andrzej.

APPLICANT: Bizindauskas, Christine B.

APPLICANT: Brauer, Andrew W.

TITLE OF INVENTION: Allergenic Proteins and

TITLE OF INVENTION: Peptides from Dog

TITLE OF INVENTION: Dander and Uses Therefor

NUMBER OF SEQUENCES: 104

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD

STREET: 60 State Street, suite 510

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII-text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/466,793

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/156,549

FILING DATE: 22-NOV-1993

APPLICATION NUMBER: 07/999,712

FILING DATE: 31-Dec-92

ATTORNEY/AGENT INFORMATION:

NAME: Mandragouras, Amy E.

REGISTRATION NUMBER: 36,207

REFERENCE/DOCKET NUMBER: IMI-026CP(IPC-048CP)

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:

LENGTH: 7 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: internal

US-08-466-793-30

Query Match 42.9%; Score 3; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0;

QY 3 EXP 5
Db 5 EXP 7

RESULT 94
US-08-491-861A-30
; Sequence 30, Application US/08491861A
; Patent No. 5939283
; GENERAL INFORMATION:
; APPLICANT: Morgenstern, Jay P.
; APPLICANT: Kanieczny, Andrzej
; APPLICANT: Bizindauskas, Christine B.
; APPLICANT: Brauer, Andrew W.
; TITLE OF INVENTION: Allergenic Proteins and Peptides from Dog
; NUMBER OF SEQUENCES: 104
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII-text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/491,861A
; FILING DATE: 27-OCT-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/999,712
; FILING DATE: 31-Dec-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: IMI-026CP (IPC-048CP)
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-491-861A-30
Query Match 42.9%; Score 3; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 EXP 5
Db 5 EXP 7

RESULT 95
US-08-310-912A-82
; Sequence 82, Application US/08310912A
; Patent No. 5981730
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Staskawicz, Brian J.
; APPLICANT: Brent, Andrew F.
; APPLICANT: Dahlbeck, Douglas
; APPLICANT: Katagiri, Fumiaki
; APPLICANT: Kunkel, Barbara N.
; APPLICANT: Mindrinos, Michael N.
; APPLICANT: Yu, Guo-liang
; TITLE OF INVENTION: RPS2 GENE FAMILY, PRIMERS, PROBES, AND DETECTION
; NUMBER OF SEQUENCES: 208
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2904
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/310,912A
; FILING DATE: September 22, 1994
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/227,360
; FILING DATE: April 13, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lech, Karen F.
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/254001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 100254
; INFORMATION FOR SEQ ID NO: 82:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-310-912A-82
Query Match 42.9%; Score 3; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGE 3
Db 3 TGE 5

RESULT 96
US-09-073-663-18
; Sequence 18, Application US/09073663
; Patent No. 6127523
; GENERAL INFORMATION:
; APPLICANT: Brewton, Richard G.
; APPLICANT: Mayne, Richard
; TITLE OF INVENTION: TYPE IX COLLAGEN AND FRAGMENTS THEREOF
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McGregor & Adler, LLP
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh OS 8.1
; SOFTWARE: Microsoft Word for Macintosh
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,663
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benjamin Aaron Adler, Ph.D., J.D.
; REGISTRATION NUMBER: 35,423
; REFERENCE/DOCKET NUMBER: D5913D

TELECOMMUNICATION INFORMATION:
TELEPHONE: (713) 777-2321
TELEFAX: (713) 777-6908
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-09-073-663-18

Query Match 42.9%; Score 3; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0;

QY 2 GEK 4
DB 2 GEK 4

RESULT 97

US-08-841-089-82

; Sequence 82, Application US/08841089

; Patent No. 6127607

; GENERAL INFORMATION:

; APPLICANT: Ausubel, Frederick M.

; APPLICANT: Staskawicz, Brian J.

; APPLICANT: Brent, Andrew F.

; APPLICANT: Dahlbeck, Douglas

; APPLICANT: Katagiri, Fumiaki

; APPLICANT: Kunkel, Barbara N.

; APPLICANT: Mindrinos, Michael N.

; APPLICANT: Yu, Guo-Liang

; TITLE OF INVENTION: RPS2 GENE AND USES THEREOF

; NUMBER OF SEQUENCES: 106

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson

; STREET: 225 Franklin Street Suite 3100

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110-2904

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30B

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/841,089

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/227,360

; FILING DATE: 13-APR-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Clark, Paul T.

; REGISTRATION NUMBER: 30,162

; REFERENCE/DOCKET NUMBER: 00786/230001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 542-5070

; TELEFAX: (617) 542-8906

; INFORMATION FOR SEQ ID NO: 82:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 7 amino acids

; TYPE: amino acid

; STRANDEDNESS: not relevant

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-841-089-82

Query Match 42.9%; Score 3; DB 3; Length 7;

Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0;

QY 1 TGE 3
DB 3 TGE 5

RESULT 98

US-08-470-335-5

; Sequence 5, Application US/08470335F

; Patent No. 6147190

; GENERAL INFORMATION:

; APPLICANT: GOODEARL, ANDREW

; APPLICANT: STROOBANT, PAUL

; APPLICANT: MINGHETTI, LUISA

; APPLICANT: WATERFIELD, MICHAEL

; APPLICANT: MARCHIONNI, MARK

; APPLICANT: CHEN, MARIO S.

; APPLICANT: HILES, IAN

; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR

; TITLE OF INVENTION: PREPARATION AND USE

; FILE REFERENCE: 04585/00200B

; CURRENT APPLICATION NUMBER: US/08/470,335F

; CURRENT FILING DATE: 1995-06-06

; EARLIER APPLICATION NUMBER: 08/036,555

; EARLIER FILING DATE: 1993-03-24

; NUMBER OF SEQ ID NOS: 252

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5

; LENGTH: 7

; TYPE: PRT

; ORGANISM: Bos taurus

; FEATURE:

; NAME/KEY: VARIANT

; LOCATION: (1)...(1)

; OTHER INFORMATION: Xaa in position 1 is Lysine or Arginine

US-08-470-335-5

Query Match

42.9%; Score 3; DB 3; Length 7;

Best Local Similarity 100.0%; Pred. No. 2.5e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0;

QY 2 GEK 4
DB 3 GEK 5

RESULT 99

US-08-735-021-5

; Sequence 5, Application US/08735021B

; Patent No. 6194377

; GENERAL INFORMATION:

; APPLICANT: GOODEARL, ANDREW

; APPLICANT: STROOBANT, PAUL

; APPLICANT: MINGHETTI, LUISA

; APPLICANT: WATERFIELD, MICHAEL

; APPLICANT: MARCHIONNI, MARK

; APPLICANT: CHEN, MARIO S.

; APPLICANT: HILES, IAN

; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR

; TITLE OF INVENTION: PREPARATION AND USE

; FILE REFERENCE: 04585/00200L

; CURRENT APPLICATION NUMBER: US/08/735,021B

; CURRENT FILING DATE: 1996-10-22

; EARLIER APPLICATION NUMBER: 08/472,065

; EARLIER FILING DATE: 1995-06-06

; EARLIER APPLICATION NUMBER: 08/036,555

; EARLIER FILING DATE: 1993-03-24

; EARLIER APPLICATION NUMBER: 07/965,173

; EARLIER FILING DATE: 1992-10-23

; EARLIER APPLICATION NUMBER: 07/940,389

; EARLIER FILING DATE: 1992-09-03

EARLIER APPLICATION NUMBER: 07/907,138
EARLIER FILING DATE: 1992-06-30
EARLIER APPLICATION NUMBER: 07/863,703
EARLIER FILING DATE: 1992-04-03
NUMBER OF SEQ ID NOS: 192
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 7
TYPE: PRT
ORGANISM: Bos taurus
FEATURE:
NAME/KEY: UNSURE
LOCATION: (1)...(1)
OTHER INFORMATION: Xaa in position 1 is Lysine or Arginine.
US-08-735-021-5

Query Match 42.9%; Score 3; DB 3; Length 7;
Best Local Similarity 100.0%; Pred.No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GEK 4
DB 3 GEK 5

RESULT 100
US-08-734-664A-5
Sequence 5, Application US/08734664A
Patent No. 6204241
GENERAL INFORMATION:
APPLICANT: Goodearl, Andrew
APPLICANT: Stroobant, Paul
APPLICANT: Minghetti, Luisa
APPLICANT: Waterfield, Michael
APPLICANT: Marchionni, Mark
APPLICANT: Chen, Mario
APPLICANT: Hiles, Ian
TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
PREPARATION AND USE
NUMBER OF SEQUENCES: 187
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM Compatible Pentium
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/734,664A
FILING DATE: 22-OCT-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/249,322
FILING DATE: 26-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/036,555
FILING DATE: 24-MAR-1993
PRIOR APPLICATION DATA: 07/965,173
APPLICATION NUMBER:
FILING DATE: 23-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/940,389
FILING DATE: 03-SEP-1992
PRIOR APPLICATION DATA: 07/907,138
APPLICATION NUMBER:
FILING DATE: 30-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/863,703

FILING DATE: 03-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: UK 91 07566.3
FILING DATE: 10-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Bleker-Brady, Kristina
REGISTRATION NUMBER: 39,109
REFERENCE/DOCKET NUMBER: 04585/00200J
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 428-0200
TELEFAX: (617) 428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 7
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
FEATURE:
OTHER INFORMATION: Xaa in position 1 is Lysine or Arginine.
US-08-734-664A-5

Query Match 42.9%; Score 3; DB 3; Length 7;
Best Local Similarity 100.0%; Pred.No. 2.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GEK 4
DB 3 GEK 5

Search completed: October 21, 2003, 18:30:43
Job time : 22.4444 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:31:37 ; Search time 41.6111 Seconds
(without alignments)
28.171 Million cell updates/sec

Title: US-10-057-890A-8

Perfect score: 7

Sequence: 1 TGERPK 7

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 629382 seqs, 167460630 residues

Word size : 0

Total number of hits satisfying chosen parameters: 81216

Minimum DB seq length: 0
Maximum DB seq length: 10

Post-processing: Listing first 150 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description
1	7	100.0	7	15	US-10-057-890A-8			Sequence 8, Appli
2	6	85.7	7	10	US-09-828-648-14			Sequence 14, Appl
3	6	85.7	9	11	US-09-938-864-112			Sequence 112, App
4	6	85.7	9	11	US-09-938-864-184			Sequence 184, App
5	6	85.7	9	11	US-09-791-477-112			Sequence 112, App
6	6	85.7	9	11	US-09-791-477-184			Sequence 184, App
7	6	85.7	9	11	US-09-785-019-112			Sequence 112, App
8	6	85.7	9	11	US-09-785-019-184			Sequence 184, App
9	6	85.7	9	15	US-10-125-635A-112			Sequence 112, App
10	6	85.7	9	15	US-10-125-635A-184			Sequence 184, App
11	6	85.7	9	15	US-10-002-603-112			Sequence 112, App
12	6	85.7	9	15	US-10-002-603-184			Sequence 184, App
13	5	71.4	5	9	US-09-779-233-29			Sequence 29, Appl
14	5	71.4	5	9	US-09-989-789-2			Sequence 2, Appli
15	5	71.4	5	9	US-09-851-271A-7			Sequence 7, Appli

5	71.4	5	9	US-09-925-796-5	Sequence 5, Appli
5	71.4	5	9	US-09-941-450-5	Sequence 5, Appli
5	71.4	5	10	US-09-731-558-3	Sequence 3, Appli
5	71.4	5	10	US-09-942-087A-5	Sequence 5, Appli
5	71.4	5	10	US-09-942-090-5	Sequence 5, Appli
5	71.4	5	11	US-09-995-973-7	Sequence 7, Appli
5	71.4	5	11	US-09-846-033B-211	Sequence 211, App
5	71.4	5	11	US-09-500-700-67	Sequence 67, Appl
5	71.4	5	11	US-09-990-186-2	Sequence 2, Appli
5	71.4	5	11	US-09-897-844-5	Sequence 5, Appli
5	71.4	5	11	US-09-996-484-7	Sequence 7, Appli
5	71.4	5	11	US-09-989-994-2	Sequence 2, Appli
5	71.4	5	12	US-10-303-686A-48	Sequence 48, Appl
5	71.4	5	12	US-10-245-415B-5	Sequence 5, Appli
5	71.4	5	12	US-10-412-105-29	Sequence 29, Appl
5	71.4	5	12	US-10-412-109-29	Sequence 29, Appl
5	71.4	5	12	US-10-192-078-25	Sequence 25, Appl
5	71.4	5	12	US-10-422-934-19	Sequence 19, Appl
5	71.4	5	14	US-10-146-221-13	Sequence 13, Appl
5	71.4	5	15	US-10-006-069A-211	Sequence 211, App
5	71.4	5	15	US-10-055-713-5	Sequence 5, Appli
5	71.4	5	15	US-10-209-194-11	Sequence 11, Appl
5	71.4	5	15	US-10-147-286-33	Sequence 33, Appl
5	71.4	5	15	US-10-113-424-2	Sequence 2, Appli
5	71.4	5	15	US-10-055-711-5	Sequence 5, Appli
5	71.4	9	11	US-09-938-864-91	Sequence 91, Appl
5	71.4	9	11	US-09-791-477-91	Sequence 91, Appl
5	71.4	9	11	US-09-785-019-91	Sequence 91, Appl
5	71.4	9	14	US-10-146-221-16	Sequence 16, Appl
5	71.4	9	15	US-10-125-635A-91	Sequence 91, Appl
5	71.4	9	15	US-10-002-603-91	Sequence 91, Appl
5	71.4	9	9	US-09-851-271A-6	Sequence 6, Appli
4	57.1	4	11	US-09-995-973-6	Sequence 6, Appli
4	57.1	4	11	US-09-996-484-6	Sequence 6, Appli
4	57.1	4	12	US-10-192-078-24	Sequence 24, Appl
4	57.1	4	14	US-10-146-221-20	Sequence 20, Appl
4	57.1	4	15	US-10-198-677-56	Sequence 56, Appl
5	57.1	5	15	US-10-198-677-60	Sequence 60, Appl
4	57.1	5	14	US-10-008-355-19	Sequence 19, Appl
6	57.1	6	15	US-10-198-677-124	Sequence 124, App
4	57.1	4	7	US-10-015-979-93	Sequence 93, Appl
4	57.1	4	7	US-10-198-677-62	Sequence 62, Appl
4	57.1	9	11	US-09-938-864-66	Sequence 66, Appl
4	57.1	9	11	US-09-938-864-111	Sequence 111, App
4	57.1	9	11	US-09-791-477-66	Sequence 66, Appl
4	57.1	9	11	US-09-791-477-111	Sequence 111, App
4	57.1	9	11	US-09-785-019-66	Sequence 66, Appl
4	57.1	9	11	US-09-785-019-111	Sequence 111, App
4	57.1	9	15	US-10-125-635A-66	Sequence 66, Appl
4	57.1	9	15	US-10-125-635A-111	Sequence 111, App
4	57.1	9	15	US-10-002-603-66	Sequence 66, Appl
4	57.1	10	10	US-10-002-603-111	Sequence 111, App
4	57.1	10	10	US-09-753-126-126	Sequence 126, App
4	57.1	10	11	US-09-896-896A-90	Sequence 90, Appl
4	57.1	10	15	US-09-896-896A-90	Sequence 90, Appl
4	57.1	4	11	US-09-500-700-32	Sequence 32, Appl
3	42.9	3	4	US-10-128-581-16	Sequence 16, Appl
3	42.9	4	15	US-10-128-581-16	Sequence 16, Appl
3	42.9	5	12	US-10-240-545A-25	Sequence 25, Appl
3	42.9	5	15	US-10-055-713-16	Sequence 16, Appl
3	42.9	5	15	US-10-206-699-133	Sequence 133, App
3	42.9	5	15	US-10-055-711-24	Sequence 24, Appl
3	42.9	5	15	US-10-198-677-84	Sequence 84, Appl
6	42.9	6	9	US-09-851-271A-10	Sequence 10, Appl
6	42.9	6	10	US-09-862-179A-12	Sequence 12, Appl
6	42.9	6	10	US-09-727-963A-28	Sequence 28, Appl
3	42.9	6	10	US-09-727-963A-81	Sequence 81, Appl
3	42.9	6	10	US-09-959-745-1	Sequence 1, Appli
3	42.9	6	11	US-09-995-973-10	Sequence 10, Appl
3	42.9	6	11	US-09-996-484-10	Sequence 10, Appl
3	42.9	6	11	US-09-954-385-427	Sequence 427, App
3	42.9	6	12	US-10-200-281-18	Sequence 18, Appl
3	42.9	6	12	US-10-271-708-5	Sequence 5, Appli
3	42.9	6	12	US-10-148-687-20	Sequence 20, Appl

89 3 42.9 6 14 US-10-141-627-17
90 3 42.9 6 15 US-10-198-677-79
91 3 42.9 7 9 US-09-736-671-30
92 3 42.9 7 9 US-09-989-789-708
93 3 42.9 7 9 US-09-989-789-708
94 3 42.9 7 9 US-09-989-789-708
95 3 42.9 7 9 US-09-989-789-813
96 3 42.9 7 9 US-09-989-789-817
97 3 42.9 7 9 US-09-989-789-2617
98 3 42.9 7 9 US-09-929-315-20
99 3 42.9 7 10 US-09-234-395-188
100 3 42.9 7 10 US-09-234-395-175
101 3 42.9 7 10 US-09-933-215-20
102 3 42.9 7 10 US-09-305-928-188
103 3 42.9 7 10 US-09-305-928-275
104 3 42.9 7 10 US-09-862-179A-9
105 3 42.9 7 10 US-09-867-852-82
106 3 42.9 7 10 US-09-264-516A-212
107 3 42.9 7 10 US-09-264-516A-299
108 3 42.9 7 10 US-09-884-767A-28
109 3 42.9 7 10 US-09-884-767A-36
110 3 42.9 7 10 US-09-884-767A-47
111 3 42.9 7 11 US-09-990-186-708
112 3 42.9 7 11 US-09-990-186-793
113 3 42.9 7 11 US-09-990-186-813
114 3 42.9 7 11 US-09-990-186-817
115 3 42.9 7 11 US-09-990-186-2617
116 3 42.9 7 11 US-09-989-994-708
117 3 42.9 7 11 US-09-989-994-793
118 3 42.9 7 11 US-09-989-994-813
119 3 42.9 7 11 US-09-989-994-817
120 3 42.9 7 11 US-09-989-994-2617
121 3 42.9 7 14 US-10-006-630-10
122 3 42.9 7 15 US-10-196-107A-30
123 3 42.9 7 15 US-10-006-869-802
124 3 42.9 7 15 US-10-006-869-889
125 3 42.9 7 15 US-10-006-869-961
126 3 42.9 7 15 US-10-006-869-1034
127 3 42.9 7 15 US-10-006-869-1241
128 3 42.9 7 15 US-10-006-869-1355
129 3 42.9 7 15 US-10-006-869-1510
130 3 42.9 7 15 US-10-006-869-1525
131 3 42.9 7 15 US-10-006-869-1640
132 3 42.9 7 15 US-10-006-869-1788
133 3 42.9 7 15 US-10-006-869-1859
134 3 42.9 7 15 US-10-006-869-1936
135 3 42.9 7 15 US-10-006-869-2061
136 3 42.9 7 15 US-10-006-869-2144
137 3 42.9 7 15 US-10-006-869-2215
138 3 42.9 7 15 US-10-006-869-2291
139 3 42.9 7 15 US-10-006-869-2530
140 3 42.9 7 15 US-10-006-869-2545
141 3 42.9 7 15 US-10-006-869-2560
142 3 42.9 7 15 US-10-006-869-2563
143 3 42.9 7 15 US-10-006-869-2575
144 3 42.9 7 15 US-10-006-869-2691
145 3 42.9 7 15 US-10-006-869-2892
146 3 42.9 7 15 US-10-006-869-2922
147 3 42.9 7 15 US-10-006-869-3248
148 3 42.9 7 15 US-10-006-869-3263
149 3 42.9 7 15 US-10-006-869-3278
150 3 42.9 7 15 US-10-006-869-3497

ALIGNMENTS

RESULT 1

US-10-057-890A-8
; Sequence 8, Application US/10057890A
; Publication No. US20030044901A1
; GENERAL INFORMATION:
; APPLICANT: Coleman, Timothy

; APPLICANT: Mansfield, Brian
; TITLE OF INVENTION: Scaffold Fusion Polypeptides, Composition for Making the Same, ar
; FILE REFERENCE: of Using the Same.
; CURRENT APPLICATION NUMBER: US/10/057,890A
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: 60/265,782
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,858
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 8
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-890A-8

Query Match 100.0%; Score 7; DB 15; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKPYK 7
Db |||||

Db 1 TGEKPYK 7

RESULT 2

US-09-828-648-14
; Sequence 14, Application US/09828648
; Patent No. US200201510A1
; GENERAL INFORMATION:
; APPLICANT: Ikezu, Tsuneya
; APPLICANT: Leisman, Gary
; APPLICANT: Carlson, Kimberly A.
; APPLICANT: Gendelman, Howard E.
; TITLE OF INVENTION: Methods and Compositions for the
; TITLE OF INVENTION: Treatment of Human Immunodeficiency Virus Infection
; FILE REFERENCE: UNMC 63156
; CURRENT APPLICATION NUMBER: US/09/828,648
; CURRENT FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/246,331
; PRIOR FILING DATE: 2000-11-06
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: ZN FING
; LOCATION: (1)...(7)
; NAME/KEY: misc_feature
; LOCATION: (7)...(7)
; OTHER INFORMATION: Xaa = any amino acid
US-09-828-648-14

Query Match 85.7%; Score 6; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKPY 6
Db |||||

Db 1 TGEKPY 6

RESULT 3

US-09-938-864-112
; Sequence 112, Application US/09938864
; Publication No. US2003007267A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: McNeill, Patricia D.

; APPLICANT: Smithgall, Molly
; APPLICANT: Moulton, Gus
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Mossman, Sally
; APPLICANT: Evans, Lawrence
; APPLICANT: Spies, A. Gregory
; APPLICANT: Boydston, Jeremy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WTL
; FILE REFERENCE: 210121.465C5
; CURRENT APPLICATION NUMBER: US/09/938.864
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 413
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 112
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-938-864-112

Query Match 85.7%; Score 6; DB 11; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKPY 6
Db 2 TGEKPY 7

RESULT 4
US-09-938-864-184
; Sequence 184, Application US/09938864
; Publication No. US20030072767A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Smithgall, Molly
; APPLICANT: Moulton, Gus
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Mossman, Sally
; APPLICANT: Evans, Lawrence
; APPLICANT: Spies, A. Gregory
; APPLICANT: Boydston, Jeremy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WTL
; FILE REFERENCE: 210121.465C5
; CURRENT APPLICATION NUMBER: US/09/938.864
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 413
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 184
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-938-864-184

Query Match 85.7%; Score 6; DB 11; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKPY 6
Db 4 TGEKPY 9

RESULT 5
US-09-791-477-112
; Sequence 112, Application US/09791477
; Publication No. US20030082194A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander

; APPLICANT: Cheever, Martin A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
; FILE REFERENCE: 2077.000200
; CURRENT APPLICATION NUMBER: US/09/791.477
; CURRENT FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: 60/184,070
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 326
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 112
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-477-112

Query Match 85.7%; Score 6; DB 11; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKPY 6
Db 2 TGEKPY 7

RESULT 6
US-09-791-477-184
; Sequence 184, Application US/09791477
; Publication No. US20030082194A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Cheever, Martin A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
; FILE REFERENCE: 2077.000200
; CURRENT APPLICATION NUMBER: US/09/791.477
; CURRENT FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: 60/184,070
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 326
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 184
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-477-184

Query Match 85.7%; Score 6; DB 11; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKPY 6
Db 4 TGEKPY 9

RESULT 7
US-09-785-019-112
; Sequence 112, Application US/09785019
; Publication No. US20030082196A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Cheever, Martin A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Smithgall, Molly
; APPLICANT: Moulton, Gus
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Sleath, Paul
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WTL
; FILE REFERENCE: 210121.465C4
; CURRENT APPLICATION NUMBER: US/09/785.019
; CURRENT FILING DATE: 2001-02-15

; NUMBER OF SEQ ID NOS: 376
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 112
 ; LENGTH: 9
 ; TYPE: PRT
 ; ORGANISM: Homo sapien
 US-09-785-019-112

Query Match 85.7%; Score 6; DB 11; Length 9;
 Best Local Similarity 100.0%; Pred. No. 5.7e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKPY 6
 Db 2 TGEKPY 7

RESULT 8

US-09-785-019-184
 ; Sequence 184, Application US/09785019
 ; Publication No. US20030082196A1
 ; GENERAL INFORMATION:

; APPLICANT: Gaiger, Alexander
 ; APPLICANT: Cheever, Martin A.
 ; APPLICANT: McNeill, Patricia D.
 ; APPLICANT: Smithgall, Molly
 ; APPLICANT: Moulton, Gus
 ; APPLICANT: Vedvick, Thomas S.
 ; APPLICANT: Sleath, Paul

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
 ; FILE REFERENCE: 210121.465C4
 ; CURRENT APPLICATION NUMBER: US/09/785,019
 ; CURRENT FILING DATE: 2001-02-15

; NUMBER OF SEQ ID NOS: 376
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 184

LENGTH: 9

TYPE: PRT

ORGANISM: Homo sapien

US-09-785-019-184

Query Match 85.7%; Score 6; DB 11; Length 9;
 Best Local Similarity 100.0%; Pred. No. 5.7e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKPY 6
 Db 4 TGEKPY 9

RESULT 9

US-10-125-635A-112
 ; Sequence 112, Application US/10125635A
 ; Publication No. US20030039635A1
 ; GENERAL INFORMATION:

; APPLICANT: Gaiger, Alexander
 ; APPLICANT: Smithgall, Molly D.
 ; APPLICANT: Carter, Darrick
 ; APPLICANT: Cheever, Martin A.
 ; APPLICANT: McNeill, Patricia D.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
 ; FILE REFERENCE: 210121.465C7
 ; CURRENT APPLICATION NUMBER: US/10/125,635A
 ; CURRENT FILING DATE: 2002-07-19

NUMBER OF SEQ ID NOS: 461

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 112

LENGTH: 9

TYPE: PRT

ORGANISM: Homo sapien

US-10-125-635A-112

US-09-785-019-184

US-10-125-635A-112

Query Match 85.7%; Score 6; DB 15; Length 9;
 Best Local Similarity 100.0%; Pred. No. 5.7e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKPY 6
 Db 2 TGEKPY 7

RESULT 10

US-10-125-635A-184
 ; Sequence 184, Application US/10125635A
 ; Publication No. US20030039635A1
 ; GENERAL INFORMATION:

; APPLICANT: Gaiger, Alexander
 ; APPLICANT: Smithgall, Molly D.
 ; APPLICANT: Carter, Darrick
 ; APPLICANT: Cheever, Martin A.
 ; APPLICANT: McNeill, Patricia D.
 ; APPLICANT: Sutherland, R. Alec

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
 ; FILE REFERENCE: 210121.465C7
 ; CURRENT APPLICATION NUMBER: US/10/125,635A
 ; CURRENT FILING DATE: 2002-07-19

NUMBER OF SEQ ID NOS: 461

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 184

LENGTH: 9

TYPE: PRT

ORGANISM: Homo sapien

US-10-125-635A-184

Query Match 85.7%; Score 6; DB 15; Length 9;
 Best Local Similarity 100.0%; Pred. No. 5.7e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKPY 6
 Db 4 TGEKPY 9

RESULT 11

US-10-002-603-112
 ; Sequence 112, Application US/10002603
 ; Publication No. US2003005971A1
 ; GENERAL INFORMATION:

; APPLICANT: Gaiger, Alexander
 ; APPLICANT: McNeill, Patricia D.
 ; APPLICANT: Smithgall, Molly
 ; APPLICANT: Moulton, Gus
 ; APPLICANT: Vedvick, Thomas S.

APPLICANT: Sleath, Paul R.

APPLICANT: Mossman, Sally

APPLICANT: Evans, Lawrence

APPLICANT: Spies, A. Gregory

APPLICANT: Boydston, Jeremy

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
 ; FILE REFERENCE: 210121.465C6
 ; CURRENT APPLICATION NUMBER: US/10/002,603
 ; CURRENT FILING DATE: 2001-10-30

NUMBER OF SEQ ID NOS: 413

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 112

LENGTH: 9

TYPE: PRT

ORGANISM: Homo sapien

US-10-002-603-112

Query Match

85.7%; Score 6; DB 15; Length 9;

Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKPY 6
Db |||||
2 TGEKPY 7

RESULT 12

US-10-002-603-184
; Sequence 184, Application US/10002603
; Publication No. US20030095971A1
; GENERAL INFORMATION:
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Smithgall, Molly
; APPLICANT: Moulton, Gus
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Mossman, Sally
; APPLICANT: Evans, Lawrence
; APPLICANT: Spies, A. Gregory
; APPLICANT: Boydston, Jeremy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
; FILE REFERENCE: 210121.465C6
; CURRENT APPLICATION NUMBER: US/10/002.603
; CURRENT FILING DATE: 2001-10-30
; NUMBER OF SEQ ID NOS: 413
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 184
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-002-603-184

Query Match 85.7%; Score 6; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKPY 6
Db |||||
4 TGEKPY 9

RESULT 13

US-09-779-233-29
; Sequence 29, Application US/09779233
; Patent No. US20020045158A1
; GENERAL INFORMATION:
; APPLICANT: Case, Casey
; TITLE OF INVENTION: CELLS FOR DRUG DISCOVERY
; FILE REFERENCE: 8325-0010
; CURRENT APPLICATION NUMBER: US/09/779,233
; CURRENT FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: linker
US-09-779-233-29

Query Match 71.4%; Score 5; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5
Db |||||
1 TGEKP 5

RESULT 14

US-09-989-789-2
; Sequence 2, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / s11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: peptide linker
US-09-989-789-2

Query Match 71.4%; Score 5; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5
Db |||||
1 TGEKP 5

RESULT 15

US-09-851-271A-7
; Sequence 7, Application US/09851271A
; Patent No. US20020064824A1
; GENERAL INFORMATION:
; APPLICANT: Gendag Limited
; TITLE OF INVENTION: Screening System
; FILE REFERENCE: 674538-2003
; CURRENT APPLICATION NUMBER: US/09/851,271A
; CURRENT FILING DATE: 2001-05-08
; PRIOR APPLICATION NUMBER: PCT/GB99/03730
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: GB9824544.2
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(5)
; OTHER INFORMATION: linker
US-09-851-271A-7

Query Match 71.4%; Score 5; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5
Db |||||
1 TGEKP 5

RESULT 16

US-09-925-796-5
; Sequence 5, Application US/09925796
; Patent No. US20020081614A1
; GENERAL INFORMATION:
; APPLICANT: Case, Casey C.
; APPLICANT: Zhang, Lei

APPLICANT: Sangamo Biosciences, Inc.
TITLE OF INVENTION: Functional Genomics Using Zinc Finger Proteins
FILE REFERENCE: 019496-002000US
CURRENT APPLICATION NUMBER: US/09/925,796
CURRENT FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 09/395,448
PRIOR FILING DATE: 1999-09-14
PRIOR APPLICATION NUMBER: 09/229,037
PRIOR FILING DATE: 1999-01-12
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:linker
US-09-925-796-5

Query Match 71.4%; Score 5; DB 9; Length 5;
Best Local Similarity 100.0%; Pred.No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5
|||||
Db 1 TGEKP 5

RESULT 17
US-09-941-450-5
Sequence 5, Application US/09941450
Patent No. US20020094529A1
GENERAL INFORMATION:
APPLICANT: Case, Casey C.
APPLICANT: Urnov, Fyodor
TITLE OF INVENTION: GENE IDENTIFICATION
FILE REFERENCE: S7.US3 / 8325-0007.20
CURRENT APPLICATION NUMBER: US/09/941,450
CURRENT FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: 09/395,448
PRIOR FILING DATE: 1999-09-14
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:linker
US-09-941-450-5

Query Match 71.4%; Score 5; DB 9; Length 5;
Best Local Similarity 100.0%; Pred.No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5
|||||
Db 1 TGEKP 5

RESULT 18
US-09-731-558-3
Sequence 3, Application US/09731558
Patent No. US20020146691A1
GENERAL INFORMATION:
APPLICANT: Case, Casey Christopher
APPLICANT: Liu, Qiang
APPLICANT: Rebar, Edward J.
APPLICANT: Sangamo Biosciences, Inc.
TITLE OF INVENTION: Methods of Using Randomized Libraries of Zinc Finger
PROTEINS FOR THE IDENTIFICATION OF GENE FUNCTION
FILE REFERENCE: 019496-003210US
CURRENT APPLICATION NUMBER: US/09/731,558

CURRENT FILING DATE: 2000-12-06
PRIOR APPLICATION NUMBER: US 09/456,100
PRIOR FILING DATE: 1999-12-06
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:flexible
OTHER INFORMATION: polypeptide linker
US-09-731-558-3

Query Match 71.4%; Score 5; DB 10; Length 5;
Best Local Similarity 100.0%; Pred.No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5
|||||
Db 1 TGEKP 5

RESULT 19
US-09-942-087A-5
Sequence 5, Application US/09942087A
Patent No. US20020160940A1
GENERAL INFORMATION:
APPLICANT: Case, Casey Christopher
APPLICANT: Wolfe, Alan
APPLICANT: Urnov, Fyodor
APPLICANT: Lai, Albert
APPLICANT: Snowden, Andrew
APPLICANT: Tan, Siyuan
APPLICANT: Gregory, Philip
TITLE OF INVENTION: MODULATION OF ENDOGENOUS GENE EXPRESSION IN CELLS
FILE REFERENCE: 8325-0002.21 / S2-US5
CURRENT APPLICATION NUMBER: US/09/942,087A
CURRENT FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: 09/229,037
PRIOR FILING DATE: 1999-01-12
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:linker
US-09-942-087A-5

Query Match 71.4%; Score 5; DB 10; Length 5;
Best Local Similarity 100.0%; Pred.No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5
|||||
Db 1 TGEKP 5

RESULT 20
US-09-942-090-5
Sequence 5, Application US/09942090
Patent No. US20020164575A1
GENERAL INFORMATION:
APPLICANT: Case, Casey C.
APPLICANT: Zhang, Lei
APPLICANT: Sangamo Biosciences, Inc.
TITLE OF INVENTION: Functional Genomics Using Zinc Finger Proteins
FILE REFERENCE: 019496-002000US
CURRENT APPLICATION NUMBER: US/09/942,090
CURRENT FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: 09/395,448

; PRIOR FILING DATE: 1999-09-14
; PRIOR APPLICATION NUMBER: 09/229,007
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 09/229,037
; PRIOR FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:linker
US-09-942-090-5

Query Match 71.4%; Score 5; DB 10; Length 5;
Best Local Similarity 100.0%; Pred.No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGEKP 5
Db 1 TGEKP 5

RESULT 21
US-09-995-973-7
; Sequence 7, Application US/09995973
; Publication No. US20030024006A1
; GENERAL INFORMATION:
; APPLICANT: CHOO, Yeh
; APPLICANT: ULLMAN, Christopher G.
; TITLE OF INVENTION: GENE SWITCHES
; FILE REFERENCE: 8325-2003 / G7-US1
; CURRENT APPLICATION NUMBER: US/09/995,973
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: linker
US-09-995-973-7

Query Match 71.4%; Score 5; DB 11; Length 5;
Best Local Similarity 100.0%; Pred.No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGEKP 5
Db 1 TGEKP 5

RESULT 22
US-09-846-033B-211
; Sequence 211, Application US/09846033B
; Publication No. US2003004404A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolfe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo BioSciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; TITLE OF INVENTION: Finger Proteins
; FILE REFERENCE: 019496-005820US
; CURRENT APPLICATION NUMBER: US/09/846,033B
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: US 09/733,604

; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 211
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptide linker
US-09-846-033B-211

Query Match 71.4%; Score 5; DB 11; Length 5;
Best Local Similarity 100.0%; Pred.No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGEKP 5
Db 1 TGEKP 5

RESULT 23
US-09-500-700-67
; Sequence 67, Application US/09500700
; Publication No. US20030059767A1
; GENERAL INFORMATION:
; APPLICANT: THE SCRIPPS RESEARCH INSTITUTE
; APPLICANT: BARBAS III, Carlos F.
; APPLICANT: GOTTESFELD, Joel M.
; APPLICANT: WRIGHT, Peter E.
; TITLE OF INVENTION: ZINC FINGER PROTEIN DERIVATIVES AND METHODS THEREFOR
; FILE REFERENCE: SCRIPI160-4
; CURRENT APPLICATION NUMBER: US/09/500,700
; CURRENT FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: US 08/863,813
; PRIOR FILING DATE: 1997-05-27
; PRIOR APPLICATION NUMBER: US 08/676,318
; PRIOR FILING DATE: 1996-12-30
; PRIOR APPLICATION NUMBER: PCT/US95/00829
; PRIOR FILING DATE: 1995-01-18
; PRIOR APPLICATION NUMBER: US 08/312,604
; PRIOR FILING DATE: 1994-09-28
; PRIOR APPLICATION NUMBER: US 08/183,119
; PRIOR FILING DATE: 1994-01-18
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Linker peptide
US-09-500-700-67

Query Match 71.4%; Score 5; DB 11; Length 5;
Best Local Similarity 100.0%; Pred.No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGEKP 5
Db 1 TGEKP 5

RESULT 24
US-09-990-186-2
; Sequence 2, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; TITLE OF INVENTION: TRIPLETS BY ZINC FINGERS
; FILE REFERENCE: 8325-0011.21 / S11-US3

```
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: peptide linker
US-09-990-186-2

Query Match          71.4%; Score 5; DB 11; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y      1 TGEKP 5
      |||||
3B      1 TGEKP 5

RESULT 25
JS-09-897-844-5
Sequence 5, Application US/09897844
Publication No. US20030087817A1
GENERAL INFORMATION:
APPLICANT: Cox III, George No. US20030087817A1Albert
APPLICANT: Case, Casey Christopher
APPLICANT: Eisenberg, Stephen P.
APPLICANT: Jarvis, Eric Edward
APPLICANT: Spratt, Sharon Kaye
APPLICANT: Sangamo Biosciences, Inc.
TITLE OF INVENTION: Regulation of Endogenous Gene Expression in Cells Using
FILE REFERENCE: 8325-2004 / G8-US1
CURRENT FILING DATE: 2002-04-08
CURRENT APPLICATION NUMBER: US/09/996,484
NUMBER OF SEQ ID NOS: 64
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: linker
US-09-897-844-5

Query Match          71.4%; Score 5; DB 11; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y      1 TGEKP 5
      |||||
3B      1 TGEKP 5

RESULT 26
JS-09-996-484-7
Sequence 7, Application US/09996484
Publication No. US20030092010A1
GENERAL INFORMATION:
APPLICANT: CHOO, Yen
APPLICANT: ULLMAN, Christopher G.
TITLE OF INVENTION: MOLECULAR SWITCHES
FILE REFERENCE: 8325-2004 / G8-US1
CURRENT FILING DATE: 2002-04-08
CURRENT APPLICATION NUMBER: US/09/996,484
NUMBER OF SEQ ID NOS: 64
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Zinc Finger protein consensus linker sequence
US-10-303-686A-48
```

```
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: linker
US-09-996-484-7

Query Match          71.4%; Score 5; DB 11; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y      1 TGEKP 5
      |||||
3B      1 TGEKP 5

RESULT 27
US-09-989-994-2
Sequence 2, Application US/09989994
Publication No. US20030104526A1
GENERAL INFORMATION:
APPLICANT: LIO, Qiang
TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
FILE REFERENCE: 8325-0011.20 / S11-US2
CURRENT FILING DATE: 2001-11-20
CURRENT APPLICATION NUMBER: US/09/989,994
NUMBER OF SEQ ID NOS: 4085
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: peptide linker
US-09-989-994-2

Query Match          71.4%; Score 5; DB 11; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y      1 TGEKP 5
      |||||
3B      1 TGEKP 5

RESULT 28
US-10-303-686A-48
Sequence 48, Application US/10303686A
Publication No. US20030153044A1
GENERAL INFORMATION:
APPLICANT: Liljedahl et al.
TITLE OF INVENTION: TISSUES OR ORGANS FOR USE IN
FILE REFERENCE: BIOBANK.004CPI
CURRENT FILING DATE: 2002-11-13
CURRENT APPLICATION NUMBER: US/10/303,686A
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 48
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Zinc Finger protein consensus linker sequence
US-10-303-686A-48
```

Query Match 71.4%; Score 5; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5
|
|
|
|
|
DB 1 TGEKP 5

RESULT 29
US-10-245-415B-5
; Sequence 5, Application US/10245415B
; Publication No. US20030166141A1
; GENERAL INFORMATION:
; APPLICANT: Case, Casey Christopher
; APPLICANT: Cox III, George N.
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Liu, Qiang
; APPLICANT: Rebar, Edward J.
; TITLE OF INVENTION: REGULATION OF ENDOGENOUS GENE EXPRESSION IN CELLS
; FILE REFERENCE: 8325-0002.22 / S2-US7
; CURRENT FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: linker
US-10-245-415B-5

Query Match 71.4%; Score 5; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5
|
|
|
|
|
DB 1 TGEKP 5

RESULT 30
US-10-412-105-29
; Sequence 29, Application US/10412105
; Publication No. US20030175790A1
; GENERAL INFORMATION:
; APPLICANT: Case, Casey
; TITLE OF INVENTION: CELLS FOR DRUG DISCOVERY
; FILE REFERENCE: 8325-0010
; CURRENT APPLICATION NUMBER: US/10/412,105
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/779,233
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: linker
US-10-412-105-29

Query Match 71.4%; Score 5; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5
|
|
|
|
|
DB 1 TGEKP 5

RESULT 31
US-10-412-109-29
; Sequence 29, Application US/10412109
; Publication No. US20030180713A1
; GENERAL INFORMATION:
; APPLICANT: Case, Casey
; TITLE OF INVENTION: CELLS FOR DRUG DISCOVERY
; FILE REFERENCE: 8325-0010
; CURRENT APPLICATION NUMBER: US/10/412,109
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: US/09/779,233
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: linker
US-10-412-109-29

Query Match 71.4%; Score 5; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5
|
|
|
|
|
DB 1 TGEKP 5

RESULT 32
US-10-192-078-25
; Sequence 25, Application US/10192078
; Publication No. US20030188331A1
; GENERAL INFORMATION:
; APPLICANT: Choo, Yen
; APPLICANT: Ullman, Christopher Graeme
; APPLICANT: Chua, Nam-Hai
; APPLICANT: Sanchez, Juan Pablo
; TITLE OF INVENTION: Regulated Gene Expression in Plants
; FILE REFERENCE: 674538-2001.3
; CURRENT APPLICATION NUMBER: US/10/192,078
; CURRENT FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: 09/732,348
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: PCT/GB00/02071
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: UK 001580.0
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: UK 001578.4
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: UK 9912635.1
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: linker sequence used to join consensus zinc finger motifs
US-10-192-078-25

Query Match 71.4%; Score 5; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5
|
|
|
|
|
DB 1 TGEKP 5

```

RESULT 33
US-10-422-934-19
Sequence 19, Application US/10422934
Publication No. US20030186841A1
GENERAL INFORMATION:
APPLICANT: Barbas, Carlos F., III
APPLICANT: Kadan, Michael
APPLICANT: Beerli, Roger
TITLE OF INVENTION: LIGAND ACTIVATED TRANSCRIPTIONAL REGULATOR PROTEINS
FILE REFERENCE: 22908-1227C
CURRENT APPLICATION NUMBER: US/10/422,934
CURRENT FILING DATE: 2003-04-23
PRIORITY APPLICATION NUMBER: 09/586,625
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: 09/433,042
PRIOR FILING DATE: 1999-10-25
NUMBER OF SEQ ID NOS: 92
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 19
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Recombinant
OTHER INFORMATION: molecule
US-10-422-934-19
Query Match 71.4%; Score 5; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGEKP 5
Db 1 TGEKP 5
|||||

RESULT 34
US-10-146-221-13
Sequence 13, Application US/10146221
Publication No. US20020173006A1
GENERAL INFORMATION:
APPLICANT: Kim, Jin-Soo
APPLICANT: Pabo, Carl O.
APPLICANT: Massachusetts Institute of Technology
TITLE OF INVENTION: Poly Zinc Finger Proteins With Improved Linkers
FILE REFERENCE: 019496-002510US
CURRENT APPLICATION NUMBER: US/10/146,221
CURRENT FILING DATE: 2002-05-13
PRIOR APPLICATION NUMBER: US/09/260,629
PRIOR FILING DATE: 1999-03-01
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 13
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:canonical
OTHER INFORMATION: "TGEKP" linker
US-10-146-221-13
Query Match 71.4%; Score 5; DB 14; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGEKP 5
Db 1 TGEKP 5
|||||

RESULT 35
US-10-057-890a-8.oligo.rapb
Sequence 211, Application US/10006069A
Publication No. US20030021776A1
GENERAL INFORMATION:
APPLICANT: Rebar, Edward
APPLICANT: Jamieson, Andrew
APPLICANT: Liu, Qiang
APPLICANT: Liu, Pei-Qi
APPLICANT: Wolffe, Alar
APPLICANT: Eisenberg, Stephen P.
APPLICANT: Jarvis, Eric
APPLICANT: Sangamo Biosciences, Inc.
TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
Finger Proteins
FILE REFERENCE: 019496-005830US
CURRENT APPLICATION NUMBER: US/10/006,069A
CURRENT FILING DATE: 2001-12-17
PRIOR APPLICATION NUMBER: US 09/733,604
PRIOR FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 09/736,083
PRIOR FILING DATE: 2000-12-12
PRIOR APPLICATION NUMBER: US 09/846,033
PRIOR FILING DATE: 2001-04-30
NUMBER OF SEQ ID NOS: 252
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 211
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: peptide linker
US-10-006-069A-211
Query Match 71.4%; Score 5; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGEKP 5
Db 1 TGEKP 5
|||||

RESULT 36
US-10-055-713-5
Sequence 5, Application US/10055713
Publication No. US20030044957A1
GENERAL INFORMATION:
APPLICANT: JAMIESON, Andrew
APPLICANT: Li, Guofu
TITLE OF INVENTION: ZINC FINGER PROTEINS FOR DNA BINDING AND GENE
REGULATION IN PLANTS
FILE REFERENCE: 8325-0026 / S26-US1
CURRENT APPLICATION NUMBER: US/10/055,713
CURRENT FILING DATE: 2002-06-17
PRIOR APPLICATION NUMBER: 60/263,445
PRIOR FILING DATE: 2001-01-22
PRIOR APPLICATION NUMBER: 60/290,716
PRIOR FILING DATE: 2001-05-11
NUMBER OF SEQ ID NOS: 105
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: peptide ZFP linker
US-10-055-713-5
Query Match 71.4%; Score 5; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGEKP 5
```

Db 1 TGEKP 5

RESULT 37
US-10-209-194-11
; Sequence 11, Application US/10209194
; Publication No. US20030051264A1
; GENERAL INFORMATION:
; APPLICANT: LILJEDAHN, MONIKA
; APPLICANT: ASPLAND, SIMON ERIC
; TITLE OF INVENTION: GENETICALLY MODIFIED COWS HAVING REDUCED
; FILE REFERENCE: BIOBANK.007A
; CURRENT FILING DATE: 2002-07-29
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 60/309,222
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/367,091
; PRIOR FILING DATE: 2002-03-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Zinc Finger protein consensus linker sequence
US-10-209-194-11

Query Match 71.4%; Score 5; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5

Db 1 TGEKP 5

RESULT 38
US-10-147-286-33
; Sequence 33, Application US/10147286
; Publication No. US20030052174A1
; GENERAL INFORMATION:
; APPLICANT: Liljedahl et al.
; TITLE OF INVENTION: TISSUES OR ORGANS FOR USE IN
; FILE REFERENCE: XENOTRANSPLANTATION
; CURRENT FILING DATE: 2002-05-14
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/291,394
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/312,125
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/367,090
; PRIOR FILING DATE: 2002-03-21
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Zinc Finger protein consensus linker sequence
US-10-147-286-33

Query Match 71.4%; Score 5; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5

Db 1 TGEKP 5

RESULT 39
US-10-113-424-2
; Sequence 2, Application US/10113424
; Publication No. US20030105593A1
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Case, Casey C.
; APPLICANT: Cox III, George N.
; APPLICANT: Jamieson, Andrew
; APPLICANT: Rebar, Edward J.
; APPLICANT: Sangamo Biosciences, Inc.
; TITLE OF INVENTION: Selection of Sites for Targeting by Zinc Finger
; TITLE OF INVENTION: Proteins and Methods of Designing Zinc Finger Proteins
; FILE REFERENCE: 019496-001800US
; CURRENT FILING DATE: 2002-03-28
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US/10/113,424
; PRIOR FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:peptide linker
US-10-113-424-2

Query Match 71.4%; Score 5; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5

Db 1 TGEKP 5

RESULT 40
US-10-055-711-5
; Sequence 5, Application US/10055711
; Publication No. US20030108880A1
; GENERAL INFORMATION:
; APPLICANT: REBAR, Edward
; APPLICANT: JAMIESON, Andrew
; TITLE OF INVENTION: MODIFIED ZINC FINGER BINDING PROTEINS
; FILE REFERENCE: 8325-0025
; CURRENT FILING DATE: 2002-09-10
; PRIOR FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: peptide linker
US-10-055-711-5

Query Match 71.4%; Score 5; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5

Db 1 TGEKP 5

RESULT 41
US-09-938-864-91
; Sequence 91, Application US/09938864

Publication No. US20030072767A1
 GENERAL INFORMATION:
 APPLICANT: Gaiger, Alexander
 APPLICANT: McNeill, Patricia D.
 APPLICANT: Smithgall, Molly
 APPLICANT: Moulton, Gus
 APPLICANT: Vedwick, Thomas S.
 APPLICANT: Sleath, Paul R.
 APPLICANT: Mossman, Sally
 APPLICANT: Evans, Lawrence
 APPLICANT: Spies, A. Gregory
 APPLICANT: Boydston, Jeremy
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
 TITLE OF INVENTION: SPECIFIC IMMUNOTHERAPY
 FILE REFERENCE: 210121.465C5
 CURRENT APPLICATION NUMBER: US-09/938,864
 CURRENT FILING DATE: 2001-08-24
 NUMBER OF SEQ ID NOS: 413
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 91
 LENGTH: 9
 TYPE: PRT
 ORGANISM: Homo sapien
 JS-09-938-864-91

Query Match 71.4%; Score 5; DB 11; Length 9;
 Best Local Similarity 100.0%; Pred. No. 5.7e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0;

2Y 2 GEXPY 6
 |||||
 Db 1 GEXPY 5

RESULT 42
 JS-09-791-477-91
 Sequence 91, Application US/09791477
 Publication No. US20030082194A1
 GENERAL INFORMATION:
 APPLICANT: Gaiger, Alexander
 APPLICANT: Cheever, Martin A.
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
 TITLE OF INVENTION: AND THERAPY OF MALIGNANT MESOTHELIOMA
 FILE REFERENCE: 2077.000200
 CURRENT APPLICATION NUMBER: US/09/791,477
 CURRENT FILING DATE: 2001-02-22
 PRIOR APPLICATION NUMBER: 60/184,070
 PRIOR FILING DATE: 2000-02-22
 NUMBER OF SEQ ID NOS: 326
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 91
 LENGTH: 9
 TYPE: PRT
 ORGANISM: Homo sapiens
 JS-09-791-477-91

Query Match 71.4%; Score 5; DB 11; Length 9;
 Best Local Similarity 100.0%; Pred. No. 5.7e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0;

2Y 2 GEXPY 6
 |||||
 Db 1 GEXPY 5

RESULT 43
 JS-09-785-019-91
 Sequence 91, Application US/09785019
 Publication No. US20030082196A1
 GENERAL INFORMATION:
 APPLICANT: Gaiger, Alexander
 APPLICANT: Cheever, Martin A.
 APPLICANT: McNeill, Patricia D.

APPLICANT: Smithgall, Molly
 APPLICANT: Moulton, Gus
 APPLICANT: Vedwick, Thomas S.
 APPLICANT: Sleath, Paul
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
 TITLE OF INVENTION: SPECIFIC IMMUNOTHERAPY
 FILE REFERENCE: 210121.465C4
 CURRENT APPLICATION NUMBER: US/09/785,019
 CURRENT FILING DATE: 2001-02-15
 NUMBER OF SEQ ID NOS: 376
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 91
 LENGTH: 9
 TYPE: PRT
 ORGANISM: Homo sapien
 US-09-785-019-91

Query Match 71.4%; Score 5; DB 11; Length 9;
 Best Local Similarity 100.0%; Pred. No. 5.7e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0;

QY 2 GEXPY 6
 |||||
 Db 1 GEXPY 5

RESULT 44
 US-10-146-221-16
 Sequence 16, Application US/10146221
 Publication No. US20020173006A1
 GENERAL INFORMATION:
 APPLICANT: Kim, Jin-Soo
 APPLICANT: Pabo, Carl O.
 APPLICANT: Massachusetts Institute of Technology
 TITLE OF INVENTION: Poly Zinc Finger Proteins With Improved Linkers
 FILE REFERENCE: 019496-00251005
 CURRENT APPLICATION NUMBER: US/10/146,221
 CURRENT FILING DATE: 2002-05-13
 PRIOR APPLICATION NUMBER: US/09/260,629
 PRIOR FILING DATE: 1999-03-01
 NUMBER OF SEQ ID NOS: 24
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 16
 LENGTH: 9
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence:linker region
 OTHER INFORMATION: in Zif268
 US-10-146-221-16

Query Match 71.4%; Score 5; DB 14; Length 9;
 Best Local Similarity 100.0%; Pred. No. 5.7e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0;

QY 1 TGEKP 5
 |||||
 Db 2 TGEKP 6

RESULT 45
 US-10-125-635A-91
 Sequence 91, Application US/10125635A
 Publication No. US20030039635A1
 GENERAL INFORMATION:
 APPLICANT: Gaiger, Alexander
 APPLICANT: Smithgall, Molly D.
 APPLICANT: Carter, Darlick
 APPLICANT: Cheever, Martin A.
 APPLICANT: McNeill, Patricia D.
 APPLICANT: Sutherland, R. Alec
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
 TITLE OF INVENTION: SPECIFIC IMMUNOTHERAPY

; FILE REFERENCE: 210121.465C7
; CURRENT APPLICATION NUMBER: US/10/125.635A
; CURRENT FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 461
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 91
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-125-635A-91

Query Match 71.4%; Score 5; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GEXPY 6
DB 1 GEXPY 5

RESULT 46

US-10-002-603-91

; Sequence 91, Application US/10002603
; Publication No. US20030095971A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Smithgall, Molly
; APPLICANT: Moulton, Gus
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Mossman, Sally
; APPLICANT: Evans, Lawrence
; APPLICANT: Spies, A. Gregory
; APPLICANT: Boydston, Jeremy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
; FILE REFERENCE: 210121.465C6
; CURRENT APPLICATION NUMBER: US/10/002.603
; CURRENT FILING DATE: 2001-10-30
; NUMBER OF SEQ ID NOS: 413
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 91
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-002-603-91

Query Match 71.4%; Score 5; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GEXPY 6
DB 1 GEXPY 5

RESULT 47

US-09-851-271A-6
; Sequence 6, Application US/09851271A
; Patent No. US20020064824A1
; GENERAL INFORMATION:
; APPLICANT: Gendag Limited
; TITLE OF INVENTION: Screening System
; FILE REFERENCE: 674538-2003
; CURRENT APPLICATION NUMBER: US/09/851.271A
; CURRENT FILING DATE: 2001-05-08
; PRIOR APPLICATION NUMBER: PCT/GB99/03730
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: GB9824544.2
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 6
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(4)
; OTHER INFORMATION: linker
US-09-851-271A-6

Query Match 57.1%; Score 4; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEK 4
DB 1 TGEK 4

RESULT 48

US-09-995-973-6
; Sequence 6, Application US/09995973
; Publication No. US20030024006A1
; GENERAL INFORMATION:
; APPLICANT: CHOO, Yen
; APPLICANT: ULLMAN, Christopher G.
; TITLE OF INVENTION: GENE SWITCHES
; FILE REFERENCE: 8325-2003 / G7-US1
; CURRENT APPLICATION NUMBER: US/09/995.973
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: linker
US-09-995-973-6

Query Match 57.1%; Score 4; DB 11; Length 4;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEK 4
DB 1 TGEK 4

RESULT 49

US-09-996-484-6
; Sequence 6, Application US/09996484
; Publication No. US20030092010A1
; GENERAL INFORMATION:
; APPLICANT: CHOO, Yen
; APPLICANT: ULLMAN, Christopher G.
; TITLE OF INVENTION: MOLECULAR SWITCHES
; FILE REFERENCE: 8325-2004 / G8-US1
; CURRENT APPLICATION NUMBER: US/09/996.484
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: linker
US-09-996-484-6

Query Match 57.1%; Score 4; DB 11; Length 4;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 TGEK 4
      ||||
Db      1 TGEK 4

RESULT 50
US-10-192-078-24
; Sequence 24, Application US/10192078
; Publication No. US20030188331A1
; GENERAL INFORMATION:
; APPLICANT: Choo, Yen
; APPLICANT: Ullman, Christopher Graeme
; APPLICANT: Chua, Nam-Hai
; APPLICANT: Sanchez, Juan Pablo
; TITLE OF INVENTION: Regulated Gene Expression in Plants
; FILE REFERENCE: 674538-2001.3
; CURRENT APPLICATION NUMBER: US/10/192,078
; CURRENT FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: 09/732,348
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: PCT/GB00/02071
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: UK 001580.0
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: UK 001578.4
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: UK 9912635.1
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: linker sequence used to join consensus zinc finger motifs
US-10-192-078-24

Query Match      57.1%; Score 4; DB 12; Length 4;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TGEK 4
      ||||
Db      1 TGEK 4

RESULT 51
US-10-146-221-20
; Sequence 20, Application US/10146221
; Publication No. US20020173006A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Pabo, Carl O.
; APPLICANT: Massachusetts Institute of Technology
; TITLE OF INVENTION: Poly Zinc Finger Proteins With Improved Linkers
; FILE REFERENCE: 019496-002510US
; CURRENT APPLICATION NUMBER: US/10/146,221
; CURRENT FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: US/09/260,629
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-146-221-20

Query Match      57.1%; Score 4; DB 12; Length 4;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TGEK 4
      ||||
Db      1 TGEK 4

RESULT 52
US-10-198-677-56
; Sequence 56, Application US/10198677
; Publication No. US20030119023A1
; GENERAL INFORMATION:
; APPLICANT: Choo, Yen
; APPLICANT: KLUG, Aaron
; APPLICANT: MOORE, Michael
; TITLE OF INVENTION: NUCLEIC ACID BINDING POLYPEPTIDES CHARACTERIZED BY
; TITLE OF INVENTION: FLEXIBLE LINKERS CONNECTED NUCLEIC ACID BINDING
; TITLE OF INVENTION: MOLECULES
; FILE REFERENCE: 8325-2011 / G11-US1
; CURRENT APPLICATION NUMBER: US/10/198,677
; CURRENT FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 56
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: canonical linker
US-10-198-677-56

Query Match      57.1%; Score 4; DB 15; Length 4;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 GEKP 5
      ||||
Db      1 GEKP 4

RESULT 53
US-10-198-677-60
; Sequence 60, Application US/10198677
; Publication No. US20030119023A1
; GENERAL INFORMATION:
; APPLICANT: Choo, Yen
; APPLICANT: KLUG, Aaron
; APPLICANT: MOORE, Michael
; TITLE OF INVENTION: NUCLEIC ACID BINDING POLYPEPTIDES CHARACTERIZED BY
; TITLE OF INVENTION: FLEXIBLE LINKERS CONNECTED NUCLEIC ACID BINDING
; TITLE OF INVENTION: MOLECULES
; FILE REFERENCE: 8325-2011 / G11-US1
; CURRENT APPLICATION NUMBER: US/10/198,677
; CURRENT FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: canonical linker
US-10-198-677-60

Query Match      57.1%; Score 4; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 GEKP 5
      ||||
Db      1 GEKP 4

RESULT 54
US-10-198-677-60
; Sequence 60, Application US/10198677
; Publication No. US20030119023A1
; GENERAL INFORMATION:
; APPLICANT: Choo, Yen
; APPLICANT: KLUG, Aaron
; APPLICANT: MOORE, Michael
; TITLE OF INVENTION: NUCLEIC ACID BINDING POLYPEPTIDES CHARACTERIZED BY
; TITLE OF INVENTION: FLEXIBLE LINKERS CONNECTED NUCLEIC ACID BINDING
; TITLE OF INVENTION: MOLECULES
; FILE REFERENCE: 8325-2011 / G11-US1
; CURRENT APPLICATION NUMBER: US/10/198,677
; CURRENT FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: canonical linker
US-10-198-677-60

Query Match      57.1%; Score 4; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 GEKP 5
      ||||
Db      1 GEKP 4
```

```

Db      .      2 GEXP 5
      ||||
RESULT 54
US-10-008-355-19
; Sequence 19, Application US/10008355
; Publication No. US20020164759A1
; GENERAL INFORMATION:
; APPLICANT: Travis, James
; APPLICANT: Potempa, Jan S
; APPLICANT: Banbula, Agnieszka
; TITLE OF INVENTION: Dipeptidylpeptidases And Methods Of Use
; FILE REFERENCE: 235.0040101
; CURRENT APPLICATION NUMBER: US/10/008,355
; CURRENT FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/246,827
; PRIOR FILING DATE: 2000-11-08
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-008-355-19

Query Match      57.1%; Score 4; DB 14; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 KPYK 7
      ||||
Db      3 KPYK 6

RESULT 55
US-10-198-677-124
; Sequence 124, Application US/10198677
; Publication No. US20030119023A1
; GENERAL INFORMATION:
; APPLICANT: CHOO, Yen
; APPLICANT: KLUG, Aaron
; APPLICANT: MOORE, Michael
; TITLE OF INVENTION: NUCLEIC ACID BINDING POLYPEPTIDES CHARACTERIZED BY
; TITLE OF INVENTION: FLEXIBLE LINKERS CONNECTED NUCLEIC ACID BINDING
; TITLE OF INVENTION: MOLECULES
; FILE REFERENCE: 8325-2011 / G11-US1
; CURRENT APPLICATION NUMBER: US/10/198,677
; CURRENT FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 124
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: linker
US-10-198-677-124

Query Match      57.1%; Score 4; DB 15; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 GEXP 5
      ||||
Db      3 GEXP 6

RESULT 56
US-10-015-979-93
; Sequence 93, Application US/10015979
; Publication No. US20030036050A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Quax, Wilhelmus J.
; APPLICANT: Verhaert, Raymond M.D.
; APPLICANT: Beekwilder, Martinus J.
; APPLICANT: Aehle, Wolfgang
; TITLE OF INVENTION: Enzyme Selection
; FILE REFERENCE: 2183-5207US
; CURRENT APPLICATION NUMBER: US/10/015,979
; CURRENT FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: PCT/NL00/00399
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/138,443
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 93
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: active CB4 mutant isolate
; OTHER INFORMATION: Y screening for starch degradation
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(7)
; OTHER INFORMATION:
US-10-015-979-93

Query Match      57.1%; Score 4; DB 15; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 GEXP 5
      ||||
Db      2 GEXP 5

RESULT 57
US-10-198-677-62
; Sequence 62, Application US/10198677
; Publication No. US20030119023A1
; GENERAL INFORMATION:
; APPLICANT: CHOO, Yen
; APPLICANT: KLUG, Aaron
; APPLICANT: MOORE, Michael
; TITLE OF INVENTION: NUCLEIC ACID BINDING POLYPEPTIDES CHARACTERIZED BY
; TITLE OF INVENTION: FLEXIBLE LINKERS CONNECTED NUCLEIC ACID BINDING
; TITLE OF INVENTION: MOLECULES
; FILE REFERENCE: 8325-2011 / G11-US1
; CURRENT APPLICATION NUMBER: US/10/198,677
; CURRENT FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 62
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: canonical linker
US-10-198-677-62

Query Match      57.1%; Score 4; DB 15; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 GEXP 5
      ||||
Db      4 GEXP 7

RESULT 58
US-09-938-864-66
; Sequence 66, Application US/09938864

```

```

; Publication No. US20030072767A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Smithgall, Molly
; APPLICANT: Moulton, Gus
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Mossman, Sally
; APPLICANT: Evans, Lawrence
; APPLICANT: Spies, A. Gregory
; APPLICANT: Boydston, Jeremy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
; TITLE OF INVENTION: SPECIFIC IMMUNOTHERAPY
; FILE REFERENCE: 210121.465C5
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 413
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 66
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapien
JS-09-938-864-66

```

```

Query Match      57.1%; Score 4; DB 11; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      3 EXPY 6
      ||||
DB      1 EXPY 4

```

```

RESULT 59
JS-09-938-864-111
; Sequence 111, Application US/09938864
; Publication No. US20030072767A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Smithgall, Molly
; APPLICANT: Moulton, Gus
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Mossman, Sally
; APPLICANT: Evans, Lawrence
; APPLICANT: Spies, A. Gregory
; APPLICANT: Boydston, Jeremy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
; TITLE OF INVENTION: SPECIFIC IMMUNOTHERAPY
; FILE REFERENCE: 210121.465C5
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 413
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 111
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapien
JS-09-938-864-111

```

```

Query Match      57.1%; Score 4; DB 11; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Y      1 TGEK 4
      ||||
b      6 TGEK 9

```

```

ESULT 60
S-09-791-477-66

```

```

; Sequence 66, Application US/09791477
; Publication No. US20030082194A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Cheever, Martin A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
; TITLE OF INVENTION: AND THERAPY OF MALIGNANT MESOTHELIOMA
; FILE REFERENCE: 2077.000200
; CURRENT APPLICATION NUMBER: US/09/791.477
; CURRENT FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: 60/184,070
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 326
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 66
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-477-66

```

```

Query Match      57.1%; Score 4; DB 11; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      3 EXPY 6
      ||||
DB      1 EXPY 4

```

```

RESULT 61
US-09-791-477-111
; Sequence 111, Application US/09791477
; Publication No. US20030082194A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Cheever, Martin A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
; TITLE OF INVENTION: AND THERAPY OF MALIGNANT MESOTHELIOMA
; FILE REFERENCE: 2077.000200
; CURRENT APPLICATION NUMBER: US/09/791.477
; CURRENT FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: 60/184,070
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 326
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 111
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-477-111

```

```

Query Match      57.1%; Score 4; DB 11; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 TGEK 4
      ||||
DB      6 TGEK 9

```

```

RESULT 62
US-09-785-019-66
; Sequence 66, Application US/09785019
; Publication No. US20030082196A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Cheever, Martin A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Smithgall, Molly
; APPLICANT: Moulton, Gus
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Sleath, Paul
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1

```

```
; TITLE OF INVENTION: SPECIFIC IMMUNOTHERAPY
; FILE REFERENCE: 210121.465C4
; CURRENT APPLICATION NUMBER: US/09/785.019
; CURRENT FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 66
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-785-019-66
```

```
Query Match          57.1%; Score 4; DB 11; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0;
```

```
QY 3 EKPX 6
   ||||
Db 1 EKPX 4
```

RESULT 63

```
US-09-785-019-111
; Sequence 111, Application US/09/785019
; Publication No. US20030082196A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Cheever, Martin A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Smithgall, Molly
; APPLICANT: Moulton, Gus
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Sleath, Paul
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
; FILE REFERENCE: 210121.465C4
; CURRENT APPLICATION NUMBER: US/09/785.019
; CURRENT FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 111
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-785-019-111
```

```
Query Match          57.1%; Score 4; DB 11; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0;
```

```
QY 1 TGEK 4
   ||||
Db 6 TGEK 9
```

RESULT 64

```
US-10-125-635A-66
; Sequence 66, Application US/10125635A
; Publication No. US20030039635A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Smithgall, Molly D.
; APPLICANT: Carter, Darrick
; APPLICANT: Cheever, Martin A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Sutherland, R. Alec
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
; FILE REFERENCE: 210121.465C7
; CURRENT APPLICATION NUMBER: US/10/125.635A
; CURRENT FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 461
; SOFTWARE: FastSEQ for Windows Version 3.0
```

```
; SEQ ID NO 66
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-125-635A-66
```

```
Query Match          57.1%; Score 4; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0;
```

```
QY 3 EKPX 6
   ||||
Db 1 EKPX 4
```

RESULT 65

```
US-10-125-635A-111
; Sequence 111, Application US/10125635A
; Publication No. US20030039635A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Smithgall, Molly D.
; APPLICANT: Carter, Darrick
; APPLICANT: Cheever, Martin A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Sutherland, R. Alec
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
; FILE REFERENCE: 210121.465C7
; CURRENT APPLICATION NUMBER: US/10/125.635A
; CURRENT FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 461
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 111
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-125-635A-111
```

```
Query Match          57.1%; Score 4; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0;
```

```
QY 1 TGEK 4
   ||||
Db 6 TGEK 9
```

RESULT 66

```
US-10-002-603-66
; Sequence 66, Application US/10002603
; Publication No. US20030095971A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Smithgall, Molly
; APPLICANT: Moulton, Gus
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Mossman, Sally
; APPLICANT: Evans, Lawrence
; APPLICANT: Spies, A. Gregory
; APPLICANT: Boydston, Jeremy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
; FILE REFERENCE: 210121.465C6
; CURRENT APPLICATION NUMBER: US/10/002.603
; CURRENT FILING DATE: 2001-10-30
; NUMBER OF SEQ ID NOS: 413
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 66
; LENGTH: 9
; TYPE: PRT
```

```
; ORGANISM: Homo sapien
US-10-002-603-66

Query Match      57.1%; Score 4; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 KPYK 6
Db      1 KPYK 4

RESULT 67
US-10-002-603-111
; Sequence 111, Application US/10002603
; Publication No. US20030095971A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Smithgall, Molly
; APPLICANT: Moulton, Gus
; APPLICANT: Vedvick, Paul R.
; APPLICANT: Mossman, Sally
; APPLICANT: Evans, Lawrence
; APPLICANT: Spies, A. Gregory
; APPLICANT: Boydston, Jeremy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
; FILE REFERENCE: 210121.465C6
; CURRENT APPLICATION NUMBER: US/10/002,603
; CURRENT FILING DATE: 2001-10-30
; NUMBER OF SEQ ID NOS: 413
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 111
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-002-603-111

Query Match      57.1%; Score 4; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TGEK 4
Db      6 TGEK 9

RESULT 68
US-09-753-126-126
; Sequence 126, Application US/09753126
; Patent No. US20020127219A1
; GENERAL INFORMATION:
; APPLICANT: OKKELS, JENS SIGURD
; APPLICANT: JENSEN, ANNE DAM
; APPLICANT: HALKIER, TORBEN
; APPLICANT: JENSEN, RIKKE BOLDING
; TITLE OF INVENTION: IMPROVED LYSOSOMAL ENZYMES AND LYSOSOMAL ENZYME
; TITLE OF INVENTION: ACTIVATORS
; FILE REFERENCE: 31-000600US
; CURRENT APPLICATION NUMBER: US/09/753,126
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: PA 1999 01891
; PRIOR FILING DATE: 1999-12-30
; PRIOR FILING DATE: 1999-12-30
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: PA 200 00865
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/210,984
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: 60/211,124
; PRIOR FILING DATE: 2000-06-12

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-896-896A-90

Query Match      57.1%; Score 4; DB 11; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      4 KPYK 7
Db      2 KPYK 5

RESULT 69
US-09-896-896A-90
; Sequence 90, Application US/09896896A
; Publication No. US20030036181A1
; GENERAL INFORMATION:
; APPLICANT: MAXYGEN APS
; TITLE OF INVENTION: PEPTIDE EXTENDED GLYCOSYLATED POLYPEPTIDES
; FILE REFERENCE: 0217us210
; CURRENT APPLICATION NUMBER: US/09/896,896A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/217,497
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: US 60/225,558
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: DK PA 2000 01027
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DK PA 2000 01092
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: PCT/DK00/00743
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: PCT/DK01/00090
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 90
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-896-896A-90

Query Match      57.1%; Score 4; DB 11; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      4 KPYK 7
Db      2 KPYK 5

RESULT 70
US-10-198-677-64
; Sequence 64, Application US/10198677
; Publication No. US20030119023A1
; GENERAL INFORMATION:
; APPLICANT: CHOO, Yen
; APPLICANT: KLUG, Aaron
```

```

; APPLICANT: MOORE, Michael
; TITLE OF INVENTION: NUCLEIC ACID BINDING POLYPEPTIDES CHARACTERIZED BY
; TITLE OF INVENTION: FLEXIBLE LINKERS CONNECTED NUCLEIC ACID BINDING
; TITLE OF INVENTION: MOLECULES
; FILE REFERENCE: 8325-2011 / G11-US1
; CURRENT APPLICATION NUMBER: US/10/198,677
; CURRENT FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 64
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: canonical linker
; OTHER INFORMATION: sequence variant
US-10-198-677-64

Query Match          57.1%; Score 4; DB 15; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GEXP 5
Db 7 GEXP 10

RESULT 71
US-09-500-700-32
; Sequence 32, Application US/09500700
; Publication No. US20030059767A1
; GENERAL INFORMATION:
; APPLICANT: THE SCRIPPS RESEARCH INSTITUTE
; APPLICANT: BARBAS III, Carlos F.
; APPLICANT: GOTTESFELD, Joel M.
; APPLICANT: WRIGHT, Peter E.
; TITLE OF INVENTION: ZINC FINGER PROTEIN DERIVATIVES AND METHODS THEREFOR
; FILE REFERENCE: SCRIPI160-4
; CURRENT APPLICATION NUMBER: US/09/500,700
; CURRENT FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: US 08/863,813
; PRIOR FILING DATE: 1997-05-27
; PRIOR APPLICATION NUMBER: US 08/676,318
; PRIOR FILING DATE: 1996-12-30
; PRIOR APPLICATION NUMBER: PCT/US95/00829
; PRIOR FILING DATE: 1995-01-18
; PRIOR APPLICATION NUMBER: US 08/312,604
; PRIOR FILING DATE: 1994-09-28
; PRIOR APPLICATION NUMBER: US 08/183,119
; PRIOR FILING DATE: 1994-01-18
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Linker peptide
; NAME/KEY: MISC FEATURE
; LOCATION: (4)..(4)
; OTHER INFORMATION: Xaa is Lys or Pro
US-09-500-700-32

Query Match          42.9%; Score 3; DB 11; Length 4;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
Db 1 TGE 3

```

```

RESULT 72
US-10-128-581-16
; Sequence 16, Application US/10128581
; Publication No. US20030104984A1
; GENERAL INFORMATION:
; APPLICANT: TSYBLOVA, IRENA
; TITLE OF INVENTION: HEMOGLOBIN ALPHA CHAIN PEPTIDE FRAGMENTS USEFUL
; FOR INHIBITING STEM CELL PROLIFERATION
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/128,581
; FILING DATE: 24-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/617,840
; FILING DATE: 17-JUL-2000
; APPLICATION NUMBER: US 09/005,546
; FILING DATE: 12-JAN-1998
; APPLICATION NUMBER: US 08/535,882
; FILING DATE: 28-SEP-1995
; APPLICATION NUMBER: US 08/316,424
; FILING DATE: 30-SEP-1994
; APPLICATION NUMBER: PCT/US94/03349
; FILING DATE: 29-MAR-1994
; APPLICATION NUMBER: US 08/040,942
; FILING DATE: 31-MAR-1993
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-10-128-581-16

Query Match          42.9%; Score 3; DB 15; Length 4;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KPY 6
Db 2 KPY 4

RESULT 73
US-10-240-545A-25
; Sequence 25, Application US/10240545A
; Publication No. US20030185828A1
; GENERAL INFORMATION:
; APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.
; TITLE OF INVENTION: No. US20030185828A1el aggrecanase
; FILE REFERENCE: 08959.0002
; CURRENT APPLICATION NUMBER: US/10/240,545A
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/JP01/11033
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: JP 2000-384300
; PRIOR FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 26

```

; SEQ ID NO 25
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-240-545A-25

Query Match 42.9%; Score 3; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
|||
Db 3 TGE 5

RESULT 74

US-10-055-713-16
; Sequence 16, Application US/10055713
; Publication No. US20030044957A1
; GENERAL INFORMATION:
; APPLICANT: JAMIESON, Andrew
; APPLICANT: LI, Guofu
; TITLE OF INVENTION: ZINC FINGER PROTEINS FOR DNA BINDING AND GENE
; TITLE OF INVENTION: REGULATION IN PLANTS
; FILE REFERENCE: 8325-0026 / S26-US1
; CURRENT APPLICATION NUMBER: US/10/055,713
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/263,445
; PRIOR FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 60/290,716
; PRIOR FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: helix-capping sequence
; NAME/KEY: SITE
; LOCATION: (4)
; OTHER INFORMATION: Xaa = either Lys or Arg
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (5)
; OTHER INFORMATION: Xaa = either Lys or Pro
US-10-055-713-16

Query Match 42.9%; Score 3; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
|||
Db 1 TGE 3

RESULT 75

US-10-206-699-133
; Sequence 133, Application US/10206699
; Publication No. US20030100510A1
; GENERAL INFORMATION:
; APPLICANT: Sundaramoorthy, M.
; APPLICANT: Hudson, B.
; TITLE OF INVENTION: Crystallized structure of Type IV Collagen NC1 Domain Hexamer
; FILE REFERENCE: MBHB 01-1017
; CURRENT APPLICATION NUMBER: US/10/206,699
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/308,523
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 60/351,289
; PRIOR FILING DATE: 2001-10-29

; PRIOR APPLICATION NUMBER: US 60/366,854
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 60/385,362
; PRIOR FILING DATE: 2002-06-03
; NUMBER OF SEQ ID NOS: 307
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 133
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-206-699-133

Query Match 42.9%; Score 3; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
|||
Db 2 TGE 4

RESULT 76

US-10-055-711-24
; Sequence 24, Application US/10055711
; Publication No. US20030108860A1
; GENERAL INFORMATION:
; APPLICANT: REBAR, Edward
; APPLICANT: JAMIESON, Andrew
; TITLE OF INVENTION: MODIFIED ZINC FINGER BINDING PROTEINS
; FILE REFERENCE: 8325-0025
; CURRENT APPLICATION NUMBER: US/10/055,711
; CURRENT FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: helix-capping sequence
; NAME/KEY: SITE
; LOCATION: (4)
; OTHER INFORMATION: where Xaa is 'Lys' or 'Arg'
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (5)
; OTHER INFORMATION: where Xaa is 'Lys' or 'Pro'
US-10-055-711-24

Query Match 42.9%; Score 3; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
|||
Db 1 TGE 3

RESULT 77

US-10-198-677-84
; Sequence 84, Application US/10198677
; Publication No. US20030119023A1
; GENERAL INFORMATION:
; APPLICANT: CHOO, Yen
; APPLICANT: KLUG, Aaron
; APPLICANT: MOORE, Michael
; TITLE OF INVENTION: NUCLEIC ACID BINDING POLYPEPTIDES CHARACTERIZED BY
; TITLE OF INVENTION: FLEXIBLE LINKERS CONNECTED NUCLEIC ACID BINDING
; TITLE OF INVENTION: MOLECULES
; FILE REFERENCE: 8325-2011 / G11-US1
; CURRENT APPLICATION NUMBER: US/10/198,677
; CURRENT FILING DATE: 2002-07-17


```
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 84
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: flexible linker
US-10-198-677-84

Query Match          42.9%; Score 3; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
   |||
Db 1 TGE 3

RESULT 78
US-09-851-271A-10
; Sequence 10, Application US/09851271A
; Patent No. US20020064824A1
; GENERAL INFORMATION:
; APPLICANT: Gendag Limited
; TITLE OF INVENTION: Screening System
; FILE REFERENCE: 674538-2003
; CURRENT APPLICATION NUMBER: US/09/851,271A
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/GB99/03730
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: GB9824544.2
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(6)
; OTHER INFORMATION: leader peptide
US-09-851-271A-10

Query Match          42.9%; Score 3; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EXP 5
   |||
Db 4 EXP 6

RESULT 79
US-09-862-179A-12
; Sequence 12, Application US/09862179A
; Patent No. US20020147306A1
; GENERAL INFORMATION:
; APPLICANT: Lin, Danny
; APPLICANT: Pawson, Anthony
; TITLE OF INVENTION: PEPTIDES THAT MODULATE THE INTERACTION OF B CLASS EPHRINS
; FILE REFERENCE: MTSI-P01-009
; CURRENT APPLICATION NUMBER: US/09/862,179A
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
```

```
; OTHER INFORMATION: PDZ inhibitory peptide
US-09-862-179A-12

Query Match          42.9%; Score 3; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PYK 7
   |||
Db 3 PYK 5

RESULT 80
US-09-727-963A-28
; Sequence 28, Application US/09727963A
; Patent No. US20020155106A1
; GENERAL INFORMATION:
; APPLICANT: V.I. Technologies, Inc.
; APPLICANT: Hammond, David J.
; TITLE OF INVENTION: METHOD OF IDENTIFYING A LIGAND FOR A TARGET MOLECULE
; FILE REFERENCE: 18242-505
; CURRENT APPLICATION NUMBER: US/09/727,963A
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/327,557
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence : peptide
; OTHER INFORMATION: ligand
US-09-727-963A-28

Query Match          42.9%; Score 3; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PYK 7
   |||
Db 4 PYK 6

RESULT 81
US-09-727-963A-81
; Sequence 81, Application US/09727963A
; Patent No. US20020155106A1
; GENERAL INFORMATION:
; APPLICANT: V.I. Technologies, Inc.
; APPLICANT: Hammond, David J.
; TITLE OF INVENTION: METHOD OF IDENTIFYING A LIGAND FOR A TARGET MOLECULE
; FILE REFERENCE: 18242-505
; CURRENT APPLICATION NUMBER: US/09/727,963A
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/327,557
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 81
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Peptide Ligand
; NAME/KEY: VARIANT
; LOCATION: (1)
; OTHER INFORMATION: wherein Xaa is D-phenylalanine
US-09-727-963A-81

Query Match          42.9%; Score 3; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
```

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PKY 7
|||
Db 4 PKY 6

RESULT 82
US-09-999-745-1
; Sequence 1, Application US/09999745
; Patent No. US20020157120A1
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: CIRCULARLY PERMUTED FLUORESCENT PROTEIN INDICATORS
; FILE REFERENCE: REGEN1470-1
; CURRENT APPLICATION NUMBER: US/09/999,745
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 09/316,920
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Linker moiety
; US-09-999-745-1

Query Match 42.9%; Score 3; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
|||
Db 3 TGE 5

RESULT 83
US-09-995-973-10
; Sequence 10, Application US/09995973
; Publication No. US20030024006A1
; GENERAL INFORMATION:
; APPLICANT: CHOO, Yen
; APPLICANT: ULLMAN, Christopher G.
; TITLE OF INVENTION: GENE SWITCHES
; FILE REFERENCE: 8325-2003 / G7-US1
; CURRENT APPLICATION NUMBER: US/09/995,973
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: leader
; OTHER INFORMATION: peptide
; US-09-995-973-10

Query Match 42.9%; Score 3; DB 11; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EXP 5
|||
b 4 EXP 6

RESULT 84
US-09-996-484-10

; Sequence 10, Application US/09996484
; Publication No. US20030092010A1
; GENERAL INFORMATION:
; APPLICANT: CHOO, Yen
; APPLICANT: ULLMAN, Christopher G.
; TITLE OF INVENTION: MOLECULAR SWITCHES
; FILE REFERENCE: 8325-2004 / G8-US1
; CURRENT APPLICATION NUMBER: US/09/996,484
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: leader
; OTHER INFORMATION: peptide
; US-09-996-484-10

Query Match 42.9%; Score 3; DB 11; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EXP 5
|||
Db 4 EXP 6

RESULT 85
US-09-954-385-427
; Sequence 427, Application US/09954385
; Publication No. US20030100467A1
; GENERAL INFORMATION:
; APPLICANT: Ahle, Wolfgang
; APPLICANT: Baldwin, Toby L.
; APPLICANT: Van Gestel, Franciscus J.C.
; APPLICANT: Janssen, Giselle G.
; APPLICANT: Murray, Christopher J.
; APPLICANT: Wang, Huaming
; APPLICANT: Winetzkzy, Deborah S.
; TITLE OF INVENTION: Binding Phenol Oxidizing Enzyme-peptide
; TITLE OF INVENTION: Complexes
; FILE REFERENCE: GC690
; CURRENT APPLICATION NUMBER: US/09/954,385
; CURRENT FILING DATE: 2001-09-12
; NUMBER OF SEQ ID NOS: 433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 427
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: binding peptide
; US-09-954-385-427

Query Match 42.9%; Score 3; DB 11; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KPY 6
|||
Db 3 KPY 5

RESULT 86
US-10-200-281-18
; Sequence 18, Application US/10200281
; Publication No. US20030143563A1
; GENERAL INFORMATION:
; APPLICANT: STRAUSS, ANDREAS
; APPLICANT: THUMM, GUNTHER
; APPLICANT: FOHLNER, JOHANNES

```

; APPLICANT: GOTZ, FRIEDRICH
; TITLE OF INVENTION: METHOD FOR IDENTIFYING A NUCLEIC ACID
; FILE REFERENCE: 10496/P65266US0
; CURRENT APPLICATION NUMBER: US/10/200,281
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: PCT/EP98/06136
; PRIOR FILING DATE: 1998-09-26
; PRIOR APPLICATION NUMBER: 97 116 841.4
; PRIOR FILING DATE: 1997-09-27
; PRIOR APPLICATION NUMBER: 97 118 755.4
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Motif
US-10-200-281-18

```

```

Query Match      42.9%; Score 3; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 TGE 3
Db      4 TGE 6

```

```

RESULT 87
US-10-271-708-5
; Sequence 5, Application US/10271708
; Publication No. US20030162200A1
; GENERAL INFORMATION:
; APPLICANT: CHOO, Yen
; APPLICANT: LIU, Xiaohai
; APPLICANT: BALASUBRAMANIAN, Shankar
; APPLICANT: PATEL, Sachin D.
; APPLICANT: ISALAN, Mark
; TITLE OF INVENTION: METHODS FOR MODULATING TELOMERASE ACTIVITY
; FILE REFERENCE: SABI-036/01US (8325-2014.01)
; CURRENT APPLICATION NUMBER: US/10/271.708
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 6
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: preferred leader peptide
US-10-271-708-5

```

```

Query Match      42.9%; Score 3; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      3 EKP 5
Db      4 EKP 6

```

```

RESULT 88
US-10-148-687-20
; Sequence 20, Application US/10148687
; Publication No. US20030185836A1
; GENERAL INFORMATION:
; APPLICANT: WINTER, Gerhard
; APPLICANT: SLADE, Martin Basil
; APPLICANT: WILLIAMS, Keith Leslie
; APPLICANT: GOOLEY, Andrew Arthur
; APPLICANT: Macquarie Research Ltd

```

```

; TITLE OF INVENTION: Cryptosporidium sporozoite antigens
; FILE REFERENCE: 047763-5019-US
; CURRENT APPLICATION NUMBER: US/10/148,687
; CURRENT FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: PCT/AU00/01492
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: AU PQ4400
; PRIOR FILING DATE: 1999-12-01
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Cryptosporidium parvum
US-10-148-687-20

```

```

Query Match      42.9%; Score 3; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 TGE 3
Db      1 TGE 3

```

```

RESULT 89
US-10-141-627-17
; Sequence 17, Application US/10141627
; Publication No. US20020176863A1
; GENERAL INFORMATION:
; APPLICANT: Dale, James B.
; TITLE OF INVENTION: ANTIGEN OF HYBRID M PROTEIN AND CARRIER
; FILE REFERENCE: 481112.404C3
; CURRENT APPLICATION NUMBER: US/10/141,627
; CURRENT FILING DATE: 2002-05-37
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hexapeptide consensus sequence
US-10-141-627-17

```

```

Query Match      42.9%; Score 3; DB 14; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 TGE 3
Db      4 TGE 6

```

```

RESULT 90
US-10-198-677-79
; Sequence 79, Application US/10198677
; Publication No. US20030119023A1
; GENERAL INFORMATION:
; APPLICANT: CHOO, Yen
; APPLICANT: KLUG, Aaron
; APPLICANT: MOORE, Michael
; TITLE OF INVENTION: NUCLEIC ACID BINDING POLYPEPTIDES CHARACTERIZED BY
; FILE REFERENCE: 8325-2011 / G11-US1
; CURRENT APPLICATION NUMBER: US/10/198,677
; CURRENT FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 79
; LENGTH: 6

```

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: leader peptide
US-10-198-677-79

Query Match 42.9%; Score 3; DB 15; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EXP 5
|||
Db 4 EXP 6

RESULT 91
US-09-374-671-30
; Sequence 30, Application US/09374671
; Patent No. US20020012963A1
; GENERAL INFORMATION:
; APPLICANT: Morgenstern, Jay P.
; APPLICANT: Kanieczny, Andrzej
; APPLICANT: Bizindauskas, Christine B.
; APPLICANT: Brauer, Andrew W.
; TITLE OF INVENTION: Allergenic Proteins and Peptides from Dog
; NUMBER OF SEQUENCES: 104
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII-text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/374,671
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/491,861
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragoras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: IMI-026CP(IPC-048CP)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal

US-09-374-671-30
Query Match 42.9%; Score 3; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3 EXP 5
|||
Db 5 EXP 7

RESULT 92
US-09-736-076-61

; Sequence 61, Application US/09736076
; Patent No. US20020049301A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson Shmuel A.
; TITLE OF INVENTION: SHORT PEPTIDES WHICH SELECTIVELY
; MODULATE THE ACTIVITY OF SERINE/THREONINE KINASES
; FILE REFERENCE: 1242.1015-009
; CURRENT APPLICATION NUMBER: US/09/736,076
; CURRENT FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 08/861,338
; PRIOR FILING DATE: 1997-05-21
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: MYRISTATE
; LOCATION: (1)...(6)
; NAME/KEY: AMIDATION
; LOCATION: (6)...(7)
; OTHER INFORMATION: cRaf
US-09-736-076-61

Query Match 42.9%; Score 3; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGE 3
|||
Db 4 TGE 6

RESULT 93
US-09-989-789-708
; Sequence 708, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 708
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example ZFP
US-09-989-789-708

Query Match 42.9%; Score 3; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGE 3
|||
Db 2 TGE 4

RESULT 94
US-09-989-789-793
; Sequence 793, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2

```
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 793
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example ZFP
US-09-989-789-793

Query Match      42.9%; Score 3; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
   |||
Db 2 TGE 4
   |||

RESULT 95
US-09-989-789-813
; Sequence 813, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDES
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 813
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example ZFP
US-09-989-789-813

Query Match      42.9%; Score 3; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
   |||
Db 2 TGE 4
   |||

RESULT 96
US-09-989-789-817
; Sequence 817, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 817
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example ZFP
US-09-989-789-817

Query Match      42.9%; Score 3; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
   |||
Db 2 TGE 4
   |||

RESULT 97
US-09-989-789-2617
; Sequence 2617, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2617
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example ZFP
US-09-989-789-2617

Query Match      42.9%; Score 3; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
   |||
Db 2 TGE 4
   |||

RESULT 98
US-09-929-315-20
; Sequence 20, Application US/09929315
; Patent No. US20020082394A1
; GENERAL INFORMATION:
; APPLICANT: Call, Katherine M.
; APPLICANT: Glaser, Thomas M.
; APPLICANT: Ito, Caryn Y.
; APPLICANT: Buckler, Alan J.
; APPLICANT: Pelletier, Jerry
; APPLICANT: Haber, Daniel A.
; APPLICANT: Rose, Elise A.
; APPLICANT: Housman, David E.
; APPLICANT: Bruening, Wendy
; APPLICANT: Darveau, Andre
; TITLE OF INVENTION: Localization and Characterization of the
; FILE REFERENCE: 0050.1312-013
; CURRENT APPLICATION NUMBER: US/09/929,315
; CURRENT FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 09/037,179
; PRIOR FILING DATE: 1998-03-09
; PRIOR APPLICATION NUMBER: US 08/102,942
; PRIOR FILING DATE: 1993-08-02
; PRIOR APPLICATION NUMBER: US 07/614,161
; PRIOR FILING DATE: 1990-11-13
; PRIOR APPLICATION NUMBER: US 07/435,780
; PRIOR FILING DATE: 1989-11-13
; PRIOR APPLICATION NUMBER: US 07/795,323
; PRIOR FILING DATE: 1991-11-20
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 7
; TYPE: PRT
```

; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: H/C link between zinc fingers
; NAME/KEY: VARIANT
; LOCATION: 4, 6, 7
; OTHER INFORMATION: xaa at position 4 can be Arg or Lys;
; OTHER INFORMATION: xaa at position 6 can be Phe or Tyr; xaa at
; OTHER INFORMATION: position 7 can be any amino acid.
US-09-929-315-20

Query Match 42.9%; Score 3; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
|||
DB 1 TGE 3

RESULT 99
US-09-234-395-188
Sequence 188, Application US/09234395
Patent No. US20020123044A1
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Byers, Stephen
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: METHODS FOR DIAGNOSING AND EVALUATING CANCER
FILE REFERENCE: 100086.407C2
CURRENT APPLICATION NUMBER: US/09/234,395
CURRENT FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 324
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 188
LENGTH: 7
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Product of
OTHER INFORMATION: Synthesis and Cyclization based on Human
OTHER INFORMATION: OB-Cadherin
FEATURE:
OTHER INFORMATION: Cyclic Peptide
US-09-234-395-188

Query Match 42.9%; Score 3; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 TGE 3
|||
b 5 TGE 7

RESULT 100
US-09-234-395-275
Sequence 275, Application US/09234395
Patent No. US20020123044A1
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Byers, Stephen
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: METHODS FOR DIAGNOSING AND EVALUATING CANCER
FILE REFERENCE: 100086.407C2
CURRENT APPLICATION NUMBER: US/09/234,395
CURRENT FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 324
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 275
LENGTH: 7
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Product of
; OTHER INFORMATION: Synthesis and Cyclization based on Human
; OTHER INFORMATION: OB-Cadherin
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide
US-09-234-395-275

Query Match 42.9%; Score 3; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
|||
DB 5 TGE 7

Search completed: October 21, 2003, 18:49:32
Job time : 45.6111 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:11:55 ; Search time 59.8889 Seconds
(without alignments)
18.552 Million cell updates/sec

Title: US-10-057-890A-8

Perfect score: 7

Sequence: 1 TGEKPYK 7

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 251420

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Listing first 150 summaries

Database : A_Geneseq_19Jun03.*

```

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	100.0	7	23	ABG32537
2	6	85.7	7	23	AAE25012
3	6	85.7	7	23	AAU79207
4	6	85.7	9	21	AAU798597
5	6	85.7	9	21	AAU98669
6	6	85.7	9	22	AAU68696
7	6	85.7	9	22	AAU68768
8	6	85.7	9	22	AAU61929
9	6	85.7	9	22	AAU62001
					Human Kruppel-type
					Human OTK18 protei
					Human Kruppel asso
					WT1 derived immuno
					WT1 derived immuno
					Human Wilm's tumou
					Human Wilm's tumou
					Human WTI immunoge
					Human WTI immunoge

10	6	85.7	9	23	ABG33166	Human WTI immunoge
11	6	85.7	9	23	ABG33238	Human WTI immunoge
12	5	71.4	5	20	AAW78384	Synthetic Cys2-his
13	5	71.4	5	20	AAW84454	Linker used in the
14	5	71.4	5	20	AAW84301	Linker used in the
15	5	71.4	5	20	AAW87789	Spacer sequence of
16	5	71.4	5	21	AAW14295	Peptide linker #1
17	5	71.4	5	21	AAW02970	Linker peptide, SE
18	5	71.4	5	22	AAW08716	Peptide linker #2
19	5	71.4	5	22	AAU05195	Zinc finger protei
20	5	71.4	5	22	AAE02026	Murine Zif268 zinc
21	5	71.4	5	22	AAE62164	Linker fragment fo
22	5	71.4	5	23	AAE28170	Linker peptide #2
23	5	71.4	5	23	AAO22206	Zinc finger protei
24	5	71.4	5	23	ABW83545	Peptide linker #1.
25	5	71.4	5	23	ABU03944	Peptide linker SEQ
26	5	71.4	5	23	ABP48192	Flexible peptide 1
27	5	71.4	5	24	ABP96165	Zinc finger protei
28	5	71.4	5	24	ABG74218	Peptide linker #1
29	5	71.4	5	24	AAE30436	Human zinc finger
30	5	71.4	6	23	ABG73580	WT1 derived immuno
31	5	71.4	9	21	AAU98576	Human Wilm's tumou
32	5	71.4	9	22	AAU68675	Human WTI immunoge
33	5	71.4	9	22	AAE61908	Human WTI immunoge
34	5	71.4	9	23	ABG33145	Leukocyte cell-der
35	5	71.4	10	17	AAU99267	Synthetic Cys2-his
36	4	57.1	4	20	AAW78383	Linker used in the
37	4	57.1	4	20	AAW84300	Zinc finger protei
38	4	57.1	4	22	AAU05194	Peptide linker for
39	4	57.1	4	22	AAE85405	Canonical linker p
40	4	57.1	4	23	ABG79360	Zinc finger protei
41	4	57.1	4	24	ABP57927	Linker peptide seq
42	4	57.1	4	24	ABP55221	Peptide linker for
43	4	57.1	5	22	ABW85409	Canonical linker p
44	4	57.1	5	23	ABG79363	Nucleic acid bindi
45	4	57.1	5	23	AAU76744	Linker amino acid
46	4	57.1	5	23	ABW05066	Zinc finger protei
47	4	57.1	5	24	ABP57930	Linker peptide seq
48	4	57.1	5	24	ABP55224	Porphyromonas ging
49	4	57.1	6	23	AAO15222	Yeast PPIase trypt
50	4	57.1	7	16	AAU72879	Bst DNA polymerase
51	4	57.1	7	16	AAU72879	Peptide linker for
52	4	57.1	7	18	AAW10599	Canonical linker p
53	4	57.1	7	22	AAW85411	Nucleic acid bindi
54	4	57.1	7	23	ABG79365	Linker amino acid
55	4	57.1	7	23	AAU76746	Linker amino acid
56	4	57.1	7	23	ABW05068	Linker peptide seq
57	4	57.1	7	24	ABP57933	Linker peptide seq
58	4	57.1	7	24	ABP55226	WT1 loop peptide K0
59	4	57.1	8	20	AAW74311	WT1 derived immuno
60	4	57.1	9	21	AAU98551	WT1 derived immuno
61	4	57.1	9	21	AAU98596	Human Wilm's tumou
62	4	57.1	9	22	AAU68650	Human Wilm's tumou
63	4	57.1	9	22	AAU68695	Human WTI immunoge
64	4	57.1	9	22	AAE61883	Human WTI immunoge
65	4	57.1	9	22	AAE61928	Bone marrow target
66	4	57.1	9	23	ABJ04613	Human WTI immunoge
67	4	57.1	9	23	ABG33120	Human WTI immunoge
68	4	57.1	9	23	ABG33155	Human cancer-relat
69	4	57.1	9	24	ABR01871	Human cancer-relat
70	4	57.1	9	24	ABR01897	Human cancer-relat
71	4	57.1	9	24	ABR02120	Human cancer-relat
72	4	57.1	9	24	ABR02125	Human cancer-relat
73	4	57.1	9	24	ABR02147	Human cancer-relat
74	4	57.1	9	24	ABR02153	Human cancer-relat
75	4	57.1	9	24	ABR02158	Human cancer-relat
76	4	57.1	9	24	ABR02271	Human cancer-relat
77	4	57.1	9	24	ABR02301	Human cancer-relat
78	4	57.1	9	24	ABR02331	Human cancer-relat
79	4	57.1	9	24	ABR02346	Human cancer-relat
80	4	57.1	9	24	ABR02470	Human cancer-relat
81	4	57.1	9	24	ABR02514	Human cancer-relat
82	4	57.1	9	24	ABR02529	Human cancer-relat

83 4 57.1 9 24 ABR02698 Human cancer-relat
84 4 57.1 9 24 ABR02699 Human cancer-relat
85 4 57.1 9 24 ABR02705 Human cancer-relat
86 4 57.1 9 24 ABR02746 Human cancer-relat
87 4 57.1 9 24 ABR02866 Human cancer-relat
88 4 57.1 9 24 ABR02889 Human cancer-relat
89 4 57.1 9 24 ABR02902 Human cancer-relat
90 4 57.1 9 24 ABR03066 Human cancer-relat
91 4 57.1 9 24 ABR03073 Human cancer-relat
92 4 57.1 9 24 ABR03127 Human cancer-relat
93 4 57.1 9 24 ABR03153 Human cancer-relat
94 4 57.1 9 24 ABR05079 Human cancer-relat
95 4 57.1 9 24 ABR05312 Human cancer-relat
96 4 57.1 9 24 ABR05327 Human cancer-relat
97 4 57.1 9 24 ABR05353 Human cancer-relat
98 4 57.1 10 22 AAU05603 N-terminal peptide
99 4 57.1 10 22 AAB85413 Peptide linker for
100 4 57.1 10 23 AAG79367 Canonical linker p
101 4 57.1 10 23 AAU76748 Nucleic acid bindi
102 4 57.1 10 23 AAU06090 Human glucocorbro
103 4 57.1 10 23 ABR05070 Linker amino acid
104 4 57.1 10 24 ABR01966 Human cancer-relat
105 4 57.1 10 24 ABR01981 Human cancer-relat
106 4 57.1 10 24 ABR02025 Human cancer-relat
107 4 57.1 10 24 ABR02054 Human cancer-relat
108 4 57.1 10 24 ABR02181 Human cancer-relat
109 4 57.1 10 24 ABR02182 Human cancer-relat
110 4 57.1 10 24 ABR02227 Human cancer-relat
111 4 57.1 10 24 ABR02252 Human cancer-relat
112 4 57.1 10 24 ABR02380 Human cancer-relat
113 4 57.1 10 24 ABR02399 Human cancer-relat
114 4 57.1 10 24 ABR02417 Human cancer-relat
115 4 57.1 10 24 ABR02433 Human cancer-relat
116 4 57.1 10 24 ABR02441 Human cancer-relat
117 4 57.1 10 24 ABR02462 Human cancer-relat
118 4 57.1 10 24 ABR02599 Human cancer-relat
119 4 57.1 10 24 ABR02614 Human cancer-relat
120 4 57.1 10 24 ABR02626 Human cancer-relat
121 4 57.1 10 24 ABR02827 Human cancer-relat
122 4 57.1 10 24 ABR02831 Human cancer-relat
123 4 57.1 10 24 ABR02848 Human cancer-relat
124 4 57.1 10 24 ABR02793 Human cancer-relat
125 4 57.1 10 24 ABR02796 Human cancer-relat
126 4 57.1 10 24 ABR02800 Human cancer-relat
127 4 57.1 10 24 ABR02809 Human cancer-relat
128 4 57.1 10 24 ABR02871 Human cancer-relat
129 4 57.1 10 24 ABR02982 Human cancer-relat
130 4 57.1 10 24 ABR02985 Human cancer-relat
131 4 57.1 10 24 ABR03016 Human cancer-relat
132 4 57.1 10 24 ABR03039 Human cancer-relat
133 4 57.1 10 24 ABR03172 Human cancer-relat
134 4 57.1 10 24 ABR03174 Human cancer-relat
135 4 57.1 10 24 ABR03195 Human cancer-relat
136 4 57.1 10 24 ABR03204 Human cancer-relat
137 4 57.1 10 24 ABR03237 Human cancer-relat
138 4 57.1 10 24 ABR05446 Zinc finger protei
139 4 57.1 10 24 ABR05493 Linker peptide seq
140 4 57.1 10 24 ABR05528 Tripeptide with an
141 3 42.9 3 22 AAG65826 Sequence of antiin
142 3 42.9 4 2 AAP10135 Oligopeptide deriv
143 3 42.9 4 15 AAR65804 Transduction domai
144 3 42.9 4 16 AAR71339 TGES motif of the
145 3 42.9 4 16 AAR71343 TGES motif of the
146 3 42.9 4 16 AAR71344 TGES motif of the
147 3 42.9 4 19 AAW61300 Peptide 15 used to
148 3 42.9 4 19 AAW61295 Peptide 10 used to
149 3 42.9 4 19 AAW61131 Peptide 10 used to
150 3 42.9 4 19 AAW61136 Peptide 15 used to

ALIGNMENTS

RESULT 1
ABG32537
ID ABG32537 standard; peptide; 7 AA.
XX
AC ABG32537;
XX
DT 15-NOV-2002 (first entry)
XX
DE Human Kruppel-type linker peptide.
XX
KW Scaffolded protein; CCR5; HIV; human immunodeficiency virus infection;
KW ECD; extracellular domain; metal chelating motif; zinc finger protein;
KW integral membrane protein; soluble loop; intracellular domain; ICD;
KW gene therapy; immunogen; viral infection; human;
KW Kruppel-type linker peptide.
XX
OS Homo sapiens.
XX
PN WO200260477-A1.
XX
PD 08-AUG-2002.
XX
PF 29-JAN-2002; 2002WO-US02377.
XX
PR 31-JAN-2001; 2001US-265782P.
XX
PR 31-JAN-2001; 2001US-265858P.
XX
FA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Coleman TA, Mansfield B;
XX
DR MPI; 2002-643357/69.
XX
PT Novel scaffolded fusion polypeptide useful for therapeutic purposes or
PT for screening molecules that bind/activate/inhibit/modulate the
PT polypeptide, comprises a functional polypeptide domain fused to a
PT scaffold domain
XX
PS Disclosure; Page 19; 64pp; English.
XX
CC The invention relates to a scaffolded fusion polypeptide comprising a
CC functional polypeptide domain fused to a scaffold domain, where the
CC functional polypeptide domain corresponds to a soluble loop of an
CC integral membrane protein (e.g. human CCR5, a transmembrane receptor
CC involved in HIV (human immunodeficiency virus) infection).
CC Also included are: (1) a polypeptide comprising a scaffold domain; (2) a
CC nucleic acid encoding the fusion polypeptide; (3) a vector cassette for
CC the expression of the fusion polypeptide comprising an expression region
CC operably linked to a promoter, where the expression region comprises a
CC number of cassettes, each of which encodes a module, domain or strand of
CC the fusion polypeptide and (4) a host cell comprising the vector or
CC nucleic acid. The fusion polypeptide is useful for screening molecules
CC that bind/activate/inhibit/modulate the fusion polypeptide, by expressing
CC the fusion polypeptide from and identifying a molecule that binds to the
CC fusion polypeptide. The fusion polypeptide is useful in diagnostic
CC methods, in assays to identify compounds that interact with loops of
CC fragments of an extracellular domain (ECD) or an intracellular domain
CC (ICD) or to rapidly assay the function of mutated portions of mutant
CC integral membrane proteins without having to produce significant
CC quantities of the entire mutant integral membrane protein, to generate
CC antibodies that recognise the integral membrane proteins from which they
CC are designed, to competitively bind the ligand of a naturally occurring
CC receptor in vitro or in vivo, to display and/or screen soluble domains
CC from protein such as integral membrane proteins, to probe the structure
CC of ECD or ICD, or both, of an integral protein membrane, to modulate the
CC activity of a receptor in vivo, and for treating or preventing viral
CC infection, preferably human HIV infection e.g. by gene therapy using the
CC encoding nucleic acid. The present sequence is a Kruppel-type linker
CC peptide suitable for inclusion in the fusion protein of the invention.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 7; DB 23; Length 7;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGEKPYK 7
| | | | |
Db 1 TGEKPYK 7

RESULT 2

AAE25012

ID AAE25012 standard; peptide; 7 AA.

XX AC

XX AC AAE25012;

XX DT

XX DT 30-OCT-2002 (first entry)

XX DE

XX DE Human OTK18 protein zinc finger consensus peptide #2.

XX KW

XX KW Human immunodeficiency virus; HIV; infection; transcriptional repressor;

XX KW

XX KW OTK18; brain; polymorphonuclear blood mononuclear cell; neuronal injury;

XX KW

XX KW CD4+ T cell; antiretroviral; mononuclear phagocyte; MP; macrophage;

XX KW

XX KW gene therapy; anti-HIV; human.

XX OS

XX OS Homo sapiens.

XX FH

XX FH Key Location/Qualifiers

XX FT

XX FT Misc-difference 7 /note= Unknown

XX FT

XX PN WO200235981-A2.

XX PN

XX PD 10-MAY-2002.

XX PD

XX PF 06-NOV-2001; 2001WO-US44336.

XX PF

XX PR 06-NOV-2000; 2000US-246331P.

XX PR

XX PR 06-APR-2001; 2001US-0828648.

XX PR

XX PA (UYNE-) UNIV NEBRASKA.

XX PA

XX PI Ikezu T, Leisman G, Carlson KA, Gendelman HE;

XX PI

XX PX WPI; 2002-519218/55.

XX PX

XX PS Example 1; Page 96; 96pp; English.

XX PS

XX CC The invention relates to methods and compositions for the treatment of human immunodeficiency virus (HIV) infection. The invention also relates to OTK18 transcriptional repressor protein and its corresponding nucleic acid. An antibody to OTK18 is useful for identifying OTK18 expression in a biological sample (e.g. polymorphonuclear blood mononuclear cells, brain tissue, macrophages and CD4+ T cells). OTK18 is used for treating HIV infection. It is useful for screening molecules that modulate or affect its activity. Its antibody is useful for identifying multinuclear giant cells in HIV encephalitic brains or immune activated mononuclear phagocytes (MP) in the brains, for fluorescent activated cell sorting (FACS) analysis of peripheral blood cells to evaluate the antiretroviral reaction of MP and for immunoprecipitating proteins from a sample containing a mixture of proteins and other biological molecules. OTK18 molecules are useful in the treatment and diagnosis of HIV infection, as research tools to identify the control of gene expression in response to HIV infection and subsequent neuronal injury. OTK18 DNA is useful in gene therapy. The present sequence is human OTK18 protein zinc finger consensus peptide used to illustrate the method of the invention.

SQ Sequence 7 AA;

Query Match

Best Local Similarity 85.7%; Score 6; DB 23; Length 7;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGEKPY 6
| | | | |
Db 1 TGEKPY 6

RESULT 3

AAU79207

ID AAU79207 standard; Peptide; 7 AA.

XX AC

XX AC AAU79207;

XX DT

XX DT 15-JUL-2002 (first entry)

XX DE

XX DE Human Kruppel associated DNA binding protein 42 related peptide.

XX KW

XX KW Human; Kruppel associated DNA binding protein 42; cytostatic;

XX KW

XX KW virucide; immunomodulator; antiinflammatory; haemostatic; tumour; cancer;

XX KW

XX KW haemopathy; HIV infection; nervous system disease; developmental disease;

XX KW

XX KW hereditary disease; endocrine system disease; immune system disease;

XX KW

XX KW immunological disease; inflammation; human immunodeficiency virus.

XX OS

XX OS Homo sapiens.

XX FH

XX FH Key Location/Qualifiers

XX FT

XX FT Misc-difference 7 /label= unknown

XX FT

XX PN WO200183541-A1.

XX PN

XX PD 08-NOV-2001.

XX PD

XX PF 28-APR-2001; 2001WO-CN00661.

XX PF

XX PR 29-APR-2000; 2000CN-0115515.

XX PR

XX PA (SHAN-) SHANGHAI BIOWINDOW GENE DEV INC.

XX PA

XX PI Mao Y, Xie Y;

XX PI

XX PX WPI; 2002-062108/08.

XX PX

XX PS Disclosure; Page 1; 38pp; Chinese.

XX PS

XX CC The invention relates to the human Kruppel associated DNA binding protein 42 and the polynucleotide encoding it. The sequences of the invention are applicable in diagnosis and treatment of different kinds of tumour, haemopathy, HIV infection, nervous system disease, developmental disease, hereditary disease, endocrine system disease, immune system disease, immunological disease and various inflammations. This sequence represents a human Kruppel associated DNA binding protein 42 related peptide.

SQ Sequence 7 AA;

Query Match

Best Local Similarity 85.7%; Score 6; DB 23; Length 7;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGEKPY 6
| | | | |
Db 1 TGEKPY 6

RESULT 4

AAU98597

ID AAU98597 standard; Peptide; 9 AA.

XX AC

AC AAY98597;
 XX
 XX 31-JUL-2000 (first entry)
 DT
 XX WTI derived immunogenic peptide SEQ ID NO:112.
 DE
 XX WTI; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
 KW metastatic disease; mouse; human; Wilm's tumour; immune response;
 KW vaccine.
 XX
 XX Homo sapiens.
 OS
 XX WO200018795-A2.
 PN
 XX 06-APR-2000.
 PD
 XX 30-SEP-1999; 99WO-US22819.
 PF
 XX 30-SEP-1998; 98US-0164223.
 PR
 XX 25-MAR-1999; 99US-0276484.
 PR
 XX (CORI-) CORIXA CORP.
 PA (GAIG/) GAIGER A.
 PA
 XX Gaiger A, Cheever M;
 PI
 XX WPI; 2000-293107/25.
 PA
 XX Novel polypeptides comprising an immunogenic portion of a native WTI
 PT polypeptide, useful for inhibiting the development of malignant
 PT diseases associated with WTI expression e.g. leukemia or cancer
 XX
 XX Claim 4; Page 160; 193pp; English.
 PS
 XX The present invention describes polypeptides (I) comprising an
 XX immunogenic portion of a native Wilm's Tumour gene product polypeptide,
 CC WTI, (or variants of the immunogenic portion retaining the ability to
 CC react with WTI-specific antisera and/or T-cell lines or clones) and
 CC comprising 16 consecutive amino acids (aa) or less of a native WTI
 CC polypeptide. The polypeptides are useful therapeutically and to
 CC manufacture medicaments for enhancing/inducing an immune response in
 CC patients. The polypeptides are useful therapeutically and to
 CC manufacture medicaments for enhancing/inducing an immune response in
 CC patients. The polypeptides, mimetics or polynucleotides can be included
 CC with a carrier/excipient in pharmaceutical compositions or with a
 CC non-specific immune response enhancer (e.g. an adjuvant or enhancer
 CC preferentially enhancing a T cell response) in vaccines. Pharmaceutical
 CC compositions and vaccines can be administered to human patients to
 CC enhance or induce an immune response specific for WTI or a cell
 CC expressing WTI, useful to inhibit the development of malignant diseases
 CC associated with WTI expression, e.g. leukemia (especially acute/chronic
 CC myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
 CC breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
 CC AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
 CC AAA13862 represent PCR primers, used in the exemplification of the
 CC present invention.
 XX
 XX Sequence 9 AA;
 SQ
 Query Match 85.7%; Score 6; DB 21; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TGEKPY 6
 DB 2 TGEKPY 7
 RESULT 5
 AAY98669
 ID AAY98669 standard; Peptide; 9 AA.
 AC
 XX AAY98669;
 XX 31-JUL-2000 (first entry)
 DT

XX WTI derived immunogenic peptide SEQ ID NO:184.
 DE
 XX WTI; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
 KW metastatic disease; mouse; human; Wilm's tumour; immune response;
 KW vaccine.
 XX
 XX Homo sapiens.
 OS
 XX WO200018795-A2.
 PN
 XX 06-APR-2000.
 PD
 XX 30-SEP-1999; 99WO-US22819.
 PF
 XX 30-SEP-1998; 98US-0164223.
 PR
 XX 25-MAR-1999; 99US-0276484.
 PR
 XX (CORI-) CORIXA CORP.
 PA (GAIG/) GAIGER A.
 PA
 XX Gaiger A, Cheever M;
 PI
 XX WPI; 2000-293107/25.
 PA
 XX Novel polypeptides comprising an immunogenic portion of a native WTI
 PT polypeptide, useful for inhibiting the development of malignant
 PT diseases associated with WTI expression e.g. leukemia or cancer
 XX
 XX Claim 4; Page 170; 193pp; English.
 PS
 XX The present invention describes polypeptides (I) comprising an
 XX immunogenic portion of a native Wilm's Tumour gene product polypeptide,
 CC WTI, (or variants of the immunogenic portion retaining the ability to
 CC react with WTI-specific antisera and/or T-cell lines or clones) and
 CC comprising 16 consecutive amino acids (aa) or less of a native WTI
 CC polypeptide. The polypeptides are useful therapeutically and to
 CC manufacture medicaments for enhancing/inducing an immune response in
 CC patients. The polypeptides, mimetics or polynucleotides can be included
 CC with a carrier/excipient in pharmaceutical compositions or with a
 CC non-specific immune response enhancer (e.g. an adjuvant or enhancer
 CC preferentially enhancing a T cell response) in vaccines. Pharmaceutical
 CC compositions and vaccines can be administered to human patients to
 CC enhance or induce an immune response specific for WTI or a cell
 CC expressing WTI, useful to inhibit the development of malignant diseases
 CC associated with WTI expression, e.g. leukemia (especially acute/chronic
 CC myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
 CC breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
 CC AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
 CC AAA13862 represent PCR primers, used in the exemplification of the
 CC present invention.
 XX
 XX Sequence 9 AA;
 SQ
 Query Match 85.7%; Score 6; DB 21; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TGEKPY 6
 DB 4 TGEKPY 9
 RESULT 6
 AAU68696
 ID AAU68696 standard; Peptide; 9 AA.
 XX
 XX AAU68696;
 AC
 XX 16-JAN-2002 (first entry)
 DT
 XX Human Wilm's tumour protein, WTI, antigenic peptide #91.
 DE
 XX

KW Human; Wilms' tumour; WT1; pleural mesothelioma; antigen;
 KW leukaemia; acute myeloid leukaemia; AML; chronic myeloid leukaemia; CML;
 KW acute lymphocytic leukaemia; ALL; myelodysplastic syndromes;
 KW myeloproliferative syndrome; cancer; cytostatic.
 XX
 OS Homo sapiens.
 XX
 PN WO200162920-A2.
 XX
 PD 30-AUG-2001.
 XX
 XX
 PF 22-FEB-2001; 2001WO-US05702.
 XX
 PR 22-FEB-2000; 2000US-184070P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Cheever MA, Gaiger A;
 XX
 DR WPI; 2001-648218/74.
 XX
 XX Composition for the treatment of mesothelioma comprises specific
 PT peptides i.e. Wilms' tumour antigen polypeptide derived antigenic
 PT fragments -
 XX
 XX Claim 1; Page 220; 242pp; English.
 XX
 CC The invention relates to the use of a composition comprising at least a
 CC first isolated peptide, of between 9 and 40 amino acids or a first
 CC nucleic acid, encoding the peptide, in the manufacture of a medicament
 CC for treating or preventing mesothelioma. The peptides are antigenic
 CC peptides derived from the Wilms' tumour protein WT1. The composition is
 CC useful for the treatment of mesothelioma, Wilms' tumour, preferably
 CC pleural mesothelioma and other WTI associated malignancies e.g.
 CC leukaemia (including acute myeloid leukaemia, ALL, and childhood ALL),
 CC leukaemia, CML, acute lymphocytic leukaemia, AML, chronic myeloid
 CC leukaemia, CML, acute lymphocytic leukaemia, ALL, and childhood ALL),
 CC myelodysplastic syndromes, myeloproliferative syndromes and cancers
 CC (e.g. breast, testicular, prostate, lung and ovarian) in mammals,
 CC preferably humans. The present sequence is an antigenic peptide of
 CC the invention derived from human WT1.
 XX
 SQ Sequence 9 AA;
 Query Match 85.7%; Score 6; DB 22; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TGEKPY 6
 Db |||||
 2 TGEKPY 7
 RESULT 7
 AAU68768
 ID AAU68768 standard; Peptide; 9 AA.
 XX
 AC AAU68768;
 XX
 DT 16-JAN-2002 (first entry)
 XX
 DE Human Wilms' tumour protein, WT1, antigenic peptide #163.
 XX
 KW Human; Wilms' tumour; WT1; pleural mesothelioma; antigen;
 KW leukaemia; acute myeloid leukaemia; AML; chronic myeloid leukaemia; CML;
 KW acute lymphocytic leukaemia; ALL; myelodysplastic syndromes;
 KW myeloproliferative syndrome; cancer; cytostatic.
 XX
 OS Homo sapiens.
 XX
 PN WO200162920-A2.
 XX
 PD 30-AUG-2001.
 XX

PF 22-FEB-2001; 2001WO-US05702.
 XX
 PR 22-FEB-2000; 2000US-184070P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Cheever MA, Gaiger A;
 XX
 DR WPI; 2001-648218/74.
 XX
 XX Composition for the treatment of mesothelioma comprises specific
 PT peptides i.e. Wilms' tumour antigen polypeptide derived antigenic
 PT fragments -
 XX
 XX Claim 1; Page 230; 242pp; English.
 XX
 CC The invention relates to the use of a composition comprising at least a
 CC first isolated peptide, of between 9 and 40 amino acids or a first
 CC nucleic acid, encoding the peptide, in the manufacture of a medicament
 CC for treating or preventing mesothelioma. The peptides are antigenic
 CC peptides derived from the Wilms' tumour protein WT1. The composition is
 CC useful for the treatment of mesothelioma, Wilms' tumour, preferably
 CC pleural mesothelioma and other WTI associated malignancies e.g.
 CC leukaemia (including acute myeloid leukaemia, ALL, and childhood ALL),
 CC leukaemia, CML, acute lymphocytic leukaemia, AML, chronic myeloid
 CC leukaemia, CML, acute lymphocytic leukaemia, ALL, and childhood ALL),
 CC myelodysplastic syndromes, myeloproliferative syndromes and cancers
 CC (e.g. breast, testicular, prostate, lung and ovarian) in mammals,
 CC preferably humans. The present sequence is an antigenic peptide of
 CC the invention derived from human WT1.
 XX
 SQ Sequence 9 AA;
 Query Match 85.7%; Score 6; DB 22; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TGEKPY 6
 Db |||||
 4 TGEKPY 9
 RESULT 8
 AAG61929
 ID AAG61929 standard; Peptide; 9 AA.
 XX
 AC AAG61929;
 XX
 DT 06-JUL-2001 (first entry)
 XX
 DE Human WT1 immunogenic peptide SEQ ID NO: 112.
 XX
 KW Human; mouse; immunotherapy; cancer; leukaemia; WT1; Wilm's tumour gene;
 KW chromosome 11p13; zinc finger transcription factor.
 XX
 OS Homo sapiens.
 XX
 PN WO200125273-A2.
 XX
 PD 12-APR-2001.
 XX
 PF 04-OCT-2000; 2000WO-US27465.
 XX
 PR 04-OCT-1999; 99US-0157459.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skeiky YAW, Xu J, Cheever MA, Reed SG;
 XX
 DR WPI; 2001-328324/34.
 XX
 PT Polypeptide comprising part of the Wilm's Tumour gene product sequence is
 PT used in the diagnosis and treatment of malignant diseases e.g. leukemia
 PT and cancer associated with WT1 -

XX PS Claim 4; Page 172; 228pp; English.

XX CC The present invention describes compositions comprising peptides derived

XX CC from the Wilm's tumour protein WT1 and methods for their use in treating

XX CC malignant diseases. Peptides derived from both the murine and human WT1

XX CC proteins are provided. The human WT1 gene is found on chromosome 11p13,

XX CC and the protein was shown to be a zinc finger transcription factor. The

XX CC immunogenic peptides of the invention are particularly useful in the

XX CC diagnosis and treatment of cancer and leukaemia. The present sequence is

XX CC a polypeptide described in the exemplification of the invention.

XX SQ Sequence 9 AA;

Query Match 85.7%; Score 6; DB 22; Length 9;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKPY 6

Db 2 TGEKPY 7

RESULT 9

AAAG62001

ID AAG62001 standard; Peptide; 9 AA.

AC AAG62001;

XX DT 06-JUL-2001 (first entry)

XX DE Human WT1 immunogenic peptide SEQ ID NO: 184.

XX KW Human; mouse; immunotherapy; cancer; leukaemia; WT1; Wilm's tumour gene;

XX KW chromosome 11p13; zinc finger transcription factor.

XX OS Homo sapiens.

XX PN WO200125273-A2.

XX PD 12-APR-2001.

XX PF 04-OCT-2000; 2000WO-US27465.

XX PR 04-OCT-1999; 99US-0157459.

XX PA (CORI-) CORIXA CORP.

XX PI Skeiky YAM, Xu J, Cheever MA, Reed SG;

XX FI WPI; 2001-328324/34.

XX DR 2001-328324/34.

XX PT Polypeptide comprising part of the Wilm's Tumour gene product sequence is

XX PT used in the diagnosis and treatment of malignant diseases e.g. leukemia

XX PT and cancer associated with WT1 -

XX PS Claim 4; Page 183; 228pp; English.

XX CC The present invention describes compositions comprising peptides derived

XX CC from the Wilm's tumour protein WT1 and methods for their use in treating

XX CC malignant diseases. Peptides derived from both the murine and human WT1

XX CC proteins are provided. The human WT1 gene is found on chromosome 11p13,

XX CC and the protein was shown to be a zinc finger transcription factor. The

XX CC immunogenic peptides of the invention are particularly useful in the

XX CC diagnosis and treatment of cancer and leukaemia. The present sequence is

XX CC a polypeptide described in the exemplification of the invention.

XX SQ Sequence 9 AA;

Query Match 85.7%; Score 6; DB 22; Length 9;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKPY 6

Db 2 TGEKPY 7

RESULT 9

AAAG62001

ID AAG62001 standard; Peptide; 9 AA.

AC AAG62001;

XX DT 06-JUL-2001 (first entry)

XX DE Human WT1 immunogenic peptide SEQ ID NO: 184.

XX KW Human; mouse; immunotherapy; cancer; leukaemia; WT1; Wilm's tumour gene;

XX KW chromosome 11p13; zinc finger transcription factor.

XX OS Homo sapiens.

XX PN WO200125273-A2.

XX PD 12-APR-2001.

XX PF 04-OCT-2000; 2000WO-US27465.

XX PR 04-OCT-1999; 99US-0157459.

XX PA (CORI-) CORIXA CORP.

XX PI Skeiky YAM, Xu J, Cheever MA, Reed SG;

XX FI WPI; 2001-328324/34.

XX DR 2001-328324/34.

XX PT Polypeptide comprising part of the Wilm's Tumour gene product sequence is

XX PT used in the diagnosis and treatment of malignant diseases e.g. leukemia

XX PT and cancer associated with WT1 -

XX PS Claim 4; Page 183; 228pp; English.

XX CC The present invention describes compositions comprising peptides derived

XX CC from the Wilm's tumour protein WT1 and methods for their use in treating

XX CC malignant diseases. Peptides derived from both the murine and human WT1

XX CC proteins are provided. The human WT1 gene is found on chromosome 11p13,

XX CC and the protein was shown to be a zinc finger transcription factor. The

XX CC immunogenic peptides of the invention are particularly useful in the

XX CC diagnosis and treatment of cancer and leukaemia. The present sequence is

XX CC a polypeptide described in the exemplification of the invention.

XX SQ Sequence 9 AA;

Query Match 85.7%; Score 6; DB 22; Length 9;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKPY 6

Db 2 TGEKPY 7

RESULT 9

AAAG62001

ID AAG62001 standard; Peptide; 9 AA.

AC AAG62001;

XX DT 06-JUL-2001 (first entry)

XX DE Human WT1 immunogenic peptide SEQ ID NO: 184.

XX KW Human; mouse; immunotherapy; cancer; leukaemia; WT1; Wilm's tumour gene;

XX KW chromosome 11p13; zinc finger transcription factor.

XX OS Homo sapiens.

XX PN WO200125273-A2.

XX PD 12-APR-2001.

XX PF 04-OCT-2000; 2000WO-US27465.

XX PR 04-OCT-1999; 99US-0157459.

XX PA (CORI-) CORIXA CORP.

XX PI Skeiky YAM, Xu J, Cheever MA, Reed SG;

XX FI WPI; 2001-328324/34.

XX DR 2001-328324/34.

XX PT Polypeptide comprising part of the Wilm's Tumour gene product sequence is

XX PT used in the diagnosis and treatment of malignant diseases e.g. leukemia

XX PT and cancer associated with WT1 -

XX PS Claim 4; Page 183; 228pp; English.

XX CC The present invention describes compositions comprising peptides derived

XX CC from the Wilm's tumour protein WT1 and methods for their use in treating

XX CC malignant diseases. Peptides derived from both the murine and human WT1

XX CC proteins are provided. The human WT1 gene is found on chromosome 11p13,

XX CC and the protein was shown to be a zinc finger transcription factor. The

XX CC immunogenic peptides of the invention are particularly useful in the

XX CC diagnosis and treatment of cancer and leukaemia. The present sequence is

XX CC a polypeptide described in the exemplification of the invention.

XX SQ Sequence 9 AA;

Query Match 85.7%; Score 6; DB 23; Length 9;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKPY 6

Db 2 TGEKPY 7

RESULT 11

ABG33238

ID ABG33238 standard; Peptide; 9 AA.

XX AC ABG33238;

XX DT 15-JUL-2002 (first entry)

XX DE Human WT1 immunogenic peptide #161.

XX KW Human; mouse; cytostatic; immunostimulant; WT1; cancer;

XX KW immune response.

QY 1 TGEKPY 6

Db 4 TGEKPY 9

RESULT 10

ABG33166

ID ABG33166 standard; Peptide; 9 AA.

XX AC ABG33166;

XX DT 15-JUL-2002 (first entry)

XX DE Human WT1 immunogenic peptide #89.

XX KW Human; mouse; cytostatic; immunostimulant; WT1; cancer;

XX KW immune response.

XX OS Homo sapiens.

XX PN WO200228414-A1.

XX PD 11-APR-2002.

XX PF 03-OCT-2001; 2001WO-US31139.

XX PR 06-OCT-2000; 2000US-0684361.

XX PR 09-OCT-2000; 2000US-0685830.

XX PR 15-FEB-2001; 2001US-0785019.

XX PR 24-AUG-2001; 2001US-0938864.

XX PA (CORI-) CORIXA CORP.

XX PA (GAIG/) GAIGER A.

XX PI Gaiger A, McNeill PD, Smithgall M, Moulton G, Vedvick TS;

XX PI Sleath PR, Mosseman S, Evans L, Spies AG, Boydston J;

XX DR WPI; 2002-352217/38.

XX PT Novel isolated WT1 polynucleotide, and encoded polypeptide, useful for

XX PT treating and diagnosing cancer in a patient -

XX PS Example 4; Page 182; 260pp; English.

XX CC The invention relates to an isolated WT1 polynucleotide (1) and

XX CC polypeptide encoded by (1). The WT1 polynucleotides and polypeptides

XX CC are used for treating and detecting cancer in a patient, and for

XX CC stimulating an immune response in patient. ABG33070-ABG33405

XX CC represent WT1 amino acid sequences of the invention.

XX SQ Sequence 9 AA;

Query Match 85.7%; Score 6; DB 23; Length 9;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKPY 6

Db 2 TGEKPY 7

RESULT 11

ABG33238

ID ABG33238 standard; Peptide; 9 AA.

XX AC ABG33238;

XX DT 15-JUL-2002 (first entry)

XX DE Human WT1 immunogenic peptide #161.

XX KW Human; mouse; cytostatic; immunostimulant; WT1; cancer;

XX KW immune response.

XX Homo sapiens.
 OS WO200228414-A1.
 PN 11-APR-2002.
 PD 03-OCT-2001; 2001WO-US31139.
 PF 06-OCT-2000; 2000US-0684361.
 PR 09-OCT-2000; 2000US-0685830.
 PR 15-FEB-2001; 2001US-0785019.
 PR 24-AUG-2001; 2001US-0938864.
 XX (CORI-) CORIXA CORP.
 PA (GAIG/) GAIGER A.
 PI Gaiger A, McNeill PD, Smithgall M, Moulton G, Vedvick TS;
 PI Sleath PR, Mossman S, Evans L, Spies AG, Boydston J;
 XX WPI; 2002-352217/38.

XX Novel isolated Wt1 polynucleotide, and encoded polypeptide, useful for
 PT treating and diagnosing cancer in a patient -
 PT Example 4; Page 193; 260pp; English.

XX The invention relates to an isolated Wt1 polynucleotide (I) and
 CC polypeptide encoded by (I). The Wt1 polynucleotides and polypeptides
 CC are used for treating and detecting cancer in a patient, and for
 CC stimulating an immune response in patient. ABG33070-ABG33405
 CC represent Wt1 amino acid sequences of the invention.

XX Sequence 9 AA;

Query Match 85.7%; Score 6; DB 23; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGEKPY 6
 DB 4 TGEKPY 9
 |||||

RESULT 12
 AAW78384
 ID AAW78384 standard; peptide; 5 AA.

XX AAW78384;
 AC 11-MAY-1999 (first entry)
 DT Synthetic Cys2-his2 zinc finger linker peptide #2.
 DE Zinc finger; target sequence; binding assay; mutant;
 XX phosphorylation site; functional domain.

XX Synthetic.
 OS WO9853060-A1.
 PN 26-NOV-1998.

XX 26-MAY-1998; 98WO-GB01516.
 XX 23-MAY-1997; 97GB-0010809.
 XX (MEDI-) MEDICAL RES COUNCIL.

XX Choo Y, Isalan M, Klug A;
 PI WPI; 1999-045309/04.

PT Rules for designing zinc finger nucleic acid binding proteins
 PT specific for any base quadruplet - relate bases in the quadruplet to
 PT specific amino acids in the alpha-helical binding motif, used to
 PT detect target nucleic acids, e.g. for identification of mutants and
 PT phosphorylation sites

XX Claim 9; Page 41; 57pp; English.

XX This sequence represent a linker peptide for a generic zinc finger of
 CC the Cys2-His2 zinc finger (ZF) class (AAW78382). The ZF are generated
 CC so that they able to bind a nucleic acid quadruplet in a target
 CC sequence, where binding to base 4 of the quadruplet by an alpha-helical
 CC ZF binding motif is determined as: (a) if base 4 is A, then position +6
 CC in the helix is Gln and position ++2 is not Asp (++) indicates a residue
 CC present in an adjacent, C-terminal ZF) and (b) if base 4 is C, then
 CC position +6 may be any residue provided ++2 is not Asp. The ZF are used
 CC to detect target nucleic acids in a binding assay, e.g. for
 CC identification of mutants (they can differentiate between single bp
 CC changes in the target) or potential phosphorylation sites, and to
 CC characterise functional domains of a protein.

XX Sequence 5 AA;

Query Match 71.4%; Score 5; DB 20; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGEKP 5
 DB 1 TGEKP 5
 |||||

RESULT 13
 AAW84454
 ID AAW84454 standard; Peptide; 5 AA.

XX AAW84454;
 AC 22-MAR-1999 (first entry)
 DT Linker used in the course of the invention.

XX zinc finger; nucleotide-binding protein; cell proliferative disorder;
 KW gene therapy; cancer; psoriasis; pemphigus vulgaris; Bechet's syndrome;
 KW lipid histiocytosis; human immune deficiency virus; HIV; viral infection;
 KW transgenic plant.

XX Synthetic.

XX WO9854311-A1.
 PN 03-DEC-1998.
 PD 27-MAY-1998; 98WO-US10801.

XX 27-MAY-1997; 97US-0863813.

XX (SCRI) SCRIPPS RES INST.

XX Barbas CF, Gottesfeld JM, Wright PE;

XX WPI; 1999-059831/05.

XX New zinc finger nucleotide-binding protein variant that modulates
 PT selected nucleotide sequence - used for treatment of proliferative
 PT and viral diseases by gene therapy, and can be made selective for
 PT any target sequence

XX Claim 7; Page 126; 158pp; English.

XX The specification describes zinc finger nucleotide-binding protein
 CC variants with at least two zinc finger modules that bind to a cellular
 CC nucleotide sequence and modulate its function. Zinc finger proteins,

CC and compositions containing them, are used to increase or reduce
 CC transcription of a gene linked to the cellular nucleotide sequence.
 CC The proteins are used specifically for treating or preventing cell
 CC proliferative disorders (in humans, animals or plants, including those
 CC induced by viruses), particularly where expressed from nucleic acid by
 CC gene therapy (including ex vivo methods). Typical diseases that can be
 CC treated are many forms of cancer, psoriasis, pemphigus vulgaris,
 CC Bechet's syndrome and lipid histiocytosis, also treatment of human
 CC immune deficiency virus (HIV) and other viral infections, and production
 CC of transgenic plants resistant to bacterial and viral diseases. The
 CC present linker is used in the course of the invention.

XX SQ Sequence 5 AA;

Query Match 71.4%; Score 5; DB 20; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5
 |||||
 Db 1 TGEKP 5

RESULT 14
 AA084301
 ID AA084301 standard; Peptide; 5 AA.

XX AC AA084301;

XX DT 18-MAR-1999 (first entry)

XX DE Linker used in the zinc fingers of the invention.

XX KW Zinc finger; nucleic acid binding protein; NABP; linker;

XX KW Cys2-His2 zinc finger; detection; gene therapy; gene delivery.

XX OS Synthetic.

XX PN WO9853059-A1.

XX PD 26-NOV-1998.

XX PF 26-MAY-1998; 98WO-GB01514.

XX PR 23-MAY-1997; 97GB-0010807.

XX PA (MEDI-) MEDICAL RES COUNCIL.

XX PI Choo Y, Isalan M, Klug A;

XX DR WPI; 1999-045308/04.

XX PT Preparation of nucleic acid binding proteins - by designing protein
 PT sequences of a Cys2-His2 zinc finger class based on a nucleic acid
 PT base triplet in a target nucleic acid sequence

XX PS Claim 9; Page 44; 62pp; English.

XX CC The present sequence represents a linker used in the zinc fingers
 CC of the invention. The specification describes a method for preparing
 CC a nucleic acid binding protein (NABP) of the Cys2-His2 zinc finger
 CC class capable of binding to a nucleic acid base triplet in a target
 CC nucleic acid sequence. Binding to the 5' base of the triplet by an
 CC alpha-helical zinc finger nucleic acid binding motif in the protein
 CC is determined as follows: (a) if the 5' base in the triplet is A,
 CC then position +6 in the alpha-helix is Glu, Asn or Val; (b) if the
 CC 5' base in the triplet is C, then position +6 in the alpha-helix is
 CC Ser, Thr, Val, Ala, Glu or Asn. The methods can be used for designing
 CC a protein which is capable of binding to any predefined nucleic acid
 CC sequence. The NABPs can be used for the detection of target nucleic
 CC acid molecules. They can also be used in gene therapy, e.g. for
 CC the delivery of functional genes into defective genes, or the delivery
 CC of nonsense nucleic acid to disrupt undesired nucleic acid.

XX SQ Sequence 5 AA;

Query Match 71.4%; Score 5; DB 20; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5
 |||||
 Db 1 TGEKP 5

RESULT 15

AA087789

ID AA087789 standard; peptide; 5 AA.

XX AC AA087789;

XX DT 09-MAR-1999 (first entry)

XX DE Spacer sequence of the zinc finger library.

XX KW Zinc finger; Cys2-His2 zinc finger class; chimeric restriction enzyme;
 fusion protein; gene therapy; nucleic acid binding protein.

XX OS Synthetic.

XX PN WO9853057-A1.

XX PD 26-NOV-1998.

XX PF 26-MAY-1998; 98WO-GB01510.

XX PR 23-MAY-1997; 97GB-0010809.

XX PA (MEDI-) MEDICAL RES COUNCIL.

XX PI Choo Y, Isalan M, Klug A;

XX WPI; 1999-024577/02.

XX PT New library of nucleic acid binding zinc finger polypeptide(s) -
 PT each polypeptide comprising more than one zinc finger which is
 PT partially randomised, useful for detecting a target nucleic acid
 PT sequence

XX PS Claim 15; Page 44; 56pp; English.

XX CC The present sequence represents a spacer sequence used in the zinc finger
 CC polypeptide library of the invention. Each polypeptide of the library
 CC comprises more than one zinc finger which has been at least partially
 CC randomised. Zinc finger proteins bind to particular nucleic acid targets.
 CC The proteins can be used for determining the presence of a target nucleic
 CC acid. The proteins of the invention can be used in the manufacture of
 CC chimeric restriction enzymes, in which a nucleic acid cleaving domain is
 CC fused to a nucleic acid binding domain comprising a zinc finger. Fusion
 CC proteins comprising a binding protein and an integrase, e.g. viral
 CC integrase, can be used to target nucleic acid sequences in vivo. In gene
 CC therapy applications, the method may be targeted to the delivery of
 CC functional genes into defective genes, or the delivery of nonsense
 CC nucleic acid. Genes may also be delivered to known repetitive stretches
 CC of nucleic acid e.g. centromeres together with an activating sequence
 CC such as an LCR. Nucleic acid binding proteins can be specifically used
 CC to knock-out cells having mutant proteins, e.g. mutant ras. They can
 CC also be used to modulate the action of transcription factors, e.g. the
 CC activity of HIV tat may be reduced by binding proteins specific for HIV
 CC TAR. The new binding proteins may also be coupled to toxic molecules,
 CC e.g. nucleases, which are capable of selectively destroying cells which
 CC comprise a mutation in their endogenous nucleic acid. The products can
 CC be used in the treatment of infections.

XX SQ Sequence 5 AA;

Query Match 71.4%; Score 5; DB 20; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5
 |||||
 DB 1 TGEKP 5

RESULT 16
 AAB14295
 ID AAB14295 standard; Peptide; 5 AA.
 XX
 AC AAB14295;
 XX
 DT 08-MAR-2001 (first entry)
 XX
 DE Peptide linker #1 used in a method for selecting ZFP target sites.
 XX
 KW Bacterial; viral; infection; oncogene; tumour suppressor gene;
 growth factor; cancer; CCR5 receptor; HIV-1; apoE; Alzheimer's disease;
 foetal haemoglobin; sickle cell anaemia; zinc finger protein.
 XX
 OS Unidentified.
 XX
 PN WO200042219-A1.
 XX
 PD 20-JUL-2000.
 XX
 PF 06-JAN-2000; 2000WO-US00388.
 XX
 PR 12-JAN-1999; 99US-0229007.
 XX
 PA (SANG-) SANGAMO BIOSCIENCES INC.
 XX
 PI Eisenberg SP, Case CC, Cox GN, Jamieson A, Rebar EJ;
 DR WPI; 2000-482840/42.
 XX
 PT Novel methods for selecting target sites for, and production of, zinc
 finger proteins, useful for controlling expression of target genes,
 e.g. for inhibiting oncogenes or treating sickle cell anemia -
 XX
 PS Disclosure; Page 14; 83pp; English.

CC Zinc finger proteins (ZFPs) bind to DNA and can modulate the expression
 of a wide range of genes. The present invention relates to a method for
 selecting a target site within a nucleic acid to be targeted by a ZFP, by
 detecting a specific 10-base motif. The present sequence is a peptide
 linker. This linker may be used to link together two or more ZFPs, to
 produce a target specificity that is the aggregate of that of the
 component ZFPs. ZFPs were designed in the present invention, which could
 bind to preselected target sites, to modulate their expression (i.e.
 activate or inhibit). The target genes may be involved in disease e.g.
 bacterial or viral genes (involved in bacterial or viral infections);
 oncogenes, tumour suppressor genes or growth factors (involved in
 cancer); CCR5 receptor gene (implicated in HIV-1 infection) or apoE gene
 (implicated in Alzheimer's disease). In addition, foetal haemoglobin
 genes may be induced by a variant ZFP for treating sickle cell anaemia.
 Furthermore, plant genes involved in herbicide or disease resistance may
 be used as target genes, to improve plant characteristics.

XX Sequence 5 AA;
 SQ

Query Match 71.4%; Score 5; DB 21; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5
 |||||
 DB 1 TGEKP 5

RESULT 17
 AAB02970
 ID AAB02970 standard; peptide; 5 AA.
 XX
 AC AAB02970;
 XX
 DT 18-SEP-2000 (first entry)
 XX
 DE Linker peptide, SEQ ID NO:111.
 XX
 KW Linker peptide; zinc finger domain; alpha helix; nucleotide binding;
 DNA binding; polydactyl protein; asymmetric target recognition;
 gene specific transcriptional regulator; gene activator; gene repressor;
 transcriptional switch; oncogene; erbB-2; cancer; tumour; gene therapy;
 transgenic animal; antiviral; anticancer; diagnosis.
 XX
 OS Synthetic.
 XX
 PN WO200023464-A2.
 XX
 PD 27-APR-2000.
 XX
 PF 14-OCT-1999; 99WO-EP07742.
 XX
 PR 16-OCT-1998; 98US-0173941.
 XX
 PA (NOVS) NOVARTIS AG.
 (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.
 (SCRI) SCRIPPS RES INST.
 XX
 PI Barbas CF;
 XX
 DR WPI; 2000-339648/29.
 XX
 PT Novel isolated and purified zinc finger nucleotide-binding proteins
 with specificity for GNN triplet sequences, useful in gene therapy and
 for regulating gene function -
 XX
 PS Claim 5; Page 19; 48pp; English.

CC The invention relates to zinc finger nucleotide-binding proteins which
 comprise 2-12, preferably 2-6, operatively linked motifs selected from
 sequences AAB02860-902875. Sequences AAB02860-902875 represent the alpha
 helical regions of zinc finger domains which specifically bind to
 target nucleotide triplets of the sequence 5'-GNN-3'. Such regions
 may be linked by the peptide linker TGEKP (AAB02970). The Cys2-His2 zinc
 finger motif is the most frequently utilised nucleic acid binding motif
 in eukaryotes, and constitutes a beta-beta-alpha fold. Nucleic acid
 recognition is achieved through specific contacts from side chains of
 amino acid residues in the alpha helix. Each zinc finger can recognise a
 sub-site of 3 bp in target DNA. Covalent linkage of multiple zinc finger
 domains allows the recognition of extended contiguous asymmetric DNA
 sequences. For example, a synthetic polydactyl protein containing six
 zinc finger domains can recognise an 18 bp sequence, and such proteins
 are potentially highly gene-specific. The novel nucleotide-binding zinc
 finger proteins may therefore be used in the development of artificial
 gene-specific transcriptional regulators. Such transcriptional switches
 may be used to regulate the expression of oncogenes such as erbB-2,
 overexpression of which is involved in malignant transformation. The
 proteins are therefore useful in the treatment of cancers, and may also
 be used to activate genes involved in fighting diseases, and to treat
 viral infections by inhibiting the synthesis of viral gene products.
 They may be used in DNA-based diagnostic applications. The proteins may
 also be used in producing functional gene knockout or activation in
 heterozygous transgenic animals. Proteins of the invention can
 discriminate between sequences which have a single base difference. This
 is manifested in a >100-fold decrease in affinity for the variant
 sequence. Gene activation and repression can be achieved by targeting
 within the gene transcript, suggesting that information obtained from
 expressed sequence tags may be sufficient for the construction of gene
 switches. The present sequence represents a specifically claimed
 linker peptide.

XX

```
SQ Sequence 5 AA;
Query Match 71.4%; Score 5; DB 21; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5
DB 1 TGEKP 5

RESULT 18
AAE08716
ID AAE08716 standard; peptide; 5 AA.
AC AAE08716;
XX
XX
XX 15-NOV-2001 (first entry)
DE Peptide linker #2 for fusion of zinc finger proteins (ZFP).
XX
XX Molecular target; zinc finger protein; ZFP; cellular process;
KW signal transduction; drug-screening.
XX
XX Synthetic.
OS
XX WO200159450-A2.
PN
XX 16-AUG-2001.
PD
XX 08-FEB-2001; 2001WO-US04301.
PF
XX 08-FEB-2000; 2000US-0181117.
PR
XX (SANG-) SANGAMO BIOSCIENCES INC.
PA
XX Case C;
PI
XX WPI; 2001-522491/57.
DR
XX Screening compound for interaction with molecular target by contacting
XX compound with cells, comprising exogenous zinc finger protein that
XX modulates expression of target, and determining values of properties of
XX cells -
XX Disclosure; Page 28; 99pp; English.
XX
XX The invention relates to a method of screening a compound for interaction
XX with a molecular target. The method involves contacting first and
XX second cells with the compound and determining the values of properties
XX of the compound. The second cell comprises an exogenous zinc finger
XX protein (ZFP) that modulates the expression of the molecular target, or
XX isolating membranes from cell comprising ZFP. The methods allow for high
XX throughput screening of candidate compound and reduces the incidence of
XX false positives. The methods are useful for screening a compound for
XX interaction with a molecular target or for screening a compound for its
XX effect on a cellular process. The method is useful for testing a compound
XX for its capacity to transduce a signal to the molecular target or its
XX capacity to block transduction of a signal through the molecular target,
XX and for performing biochemical drug-screening assays. The present
XX sequence is a peptide linker used for linking ZFPs.
XX
SQ Sequence 5 AA;
Query Match 71.4%; Score 5; DB 22; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5
DB 1 TGEKP 5

RESULT 19
AAU05195
ID AAU05195 standard; Protein; 5 AA.
XX
XX AC AAU05195;
XX
XX 24-OCT-2001 (first entry)
DT
XX Zinc finger protein binding motif linker #2.
XX
XX ZFIIIA/Zif-VPL6; sub-cellular trafficking; signal transduction;
KW diagnostic; gene therapy; agricultural technology; zinc finger.
XX
XX Synthetic.
OS
XX WO200153479-A2.
PN
XX 26-JUL-2001.
PD
XX 18-JAN-2001; 2001WO-GB00187.
PF
XX 24-JAN-2000; 2000GB-0001578.
PR 24-JAN-2000; 2000GB-0001582.
PR 30-MAY-2000; 2000WO-GB02071.
PR 30-MAY-2000; 2000WO-GB02080.
PR 07-DEC-2000; 2000GB-0029901.
XX
XX (GEND-) GENDAQ LTD.
PA
XX Choo Y, Ullman CG, Moore M;
PI
XX WPI; 2001-457607/49.
DR
XX
XX Selecting switching system components comprising first molecule, second
XX molecule and ligand, involves determining degree of binding between the
XX molecules in the presence and absence of the ligand -
XX Disclosure; Page 51; 193pp; English.
XX
XX The invention relates to a method of selecting components of a switching
XX system comprising first molecule (M1), selecting (i) a ligand (L2) which
XX is capable of modulating the interaction between a nucleic acid binding
XX molecule (NBM) and a target nucleic acid (TN). NBM, TN and the ligands
XX are useful for regulating a biological process selected from sub-cellular
XX trafficking and signal transduction, and for regulating transcription or
XX translation from a nucleic acid sequence comprising TN to which NBM binds
XX in a manner modulatable by the ligand. They are also useful for
XX modulating a biological process affecting one or more genes, such as in
XX transcription, translation, phosphorylation, replication, transport,
XX splicing, integration and recombination, in a host cell. The polypeptides
XX or ligands selected from (S) are useful for regulating a biological
XX process involving binding of first polypeptide with second polypeptide,
XX selected from sub-cellular trafficking and signal transduction. (i) is
XX also useful for screening any or all the components of the gene switch
XX system or protein switch system, and for modulating the interaction
XX between the components of the switching system. The binding molecules are
XX useful as diagnostic and research tools, and for switching or modulating
XX gene expression, especially in gene therapy applications and agricultural
XX technology applications. The ligands are also useful in the treatment of
XX diseases. Multiple target nucleic acids or polypeptides could be
XX used in a single selective step, thus enabling multiple nucleic acid or
XX polypeptide binding molecules to be isolated simultaneously, even in
XX the same physical vessel. The present sequence represents the zinc finger
XX protein binding motif linker #2 used in the method of the invention.
XX
SQ Sequence 5 AA;
Query Match 71.4%; Score 5; DB 22; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5
DB 1 TGEKP 5
```


Db. 1 TGEKP 5

RESULT 20
AAE02026
ID AAE02026 standard; peptide; 5 AA.

XX AC AAE02026;
XX DT 31-JUL-2001 (first entry)
XX DE Murine Zif268 zinc finger protein linker peptide.
XX KW Fusion protein; nucleotide-binding domain; NBD;
XX KW ligand-binding domain; LBD; transcription regulating domain; TRD;
XX KW zinc finger protein; ZFP; ligand-activated transcriptional regulator;
XX KW gene regulation; gene therapy; cell proliferative disorder; cancer;
XX KW psoriasis; pemphigus vulgaris; Behcet's syndrome; lipid histiocytosis;
XX KW Zif268; murine.
XX OS Mus sp.
XX PN WO200130843-A1.
XX PD 03-MAY-2001.
XX PF 23-OCT-2000; 2000WO-EP10430.
XX PR 25-OCT-1999; 99US-0433042.
XX PR 02-JUN-2000; 2000US-0586625.
XX PA (NOVS) NOVARTIS AG.
XX PA (SCRI) SCRIPPS RES INST.
XX PI Barbas CF, Kadan M, Beerli R;
XX DR WPI; 2001-308618/32.
XX PT New fusion protein containing nucleotide-binding and ligand-binding
XX PT domains, useful e.g. in gene therapy of cancer, provides
XX PT ligand-activated control of gene expression -
XX PS Disclosure; Page 38; 218pp; English.
XX CC The invention relates to fusion protein comprising a nucleotide-binding
XX CC domain (NBD), a ligand-binding domain (LBD) of an intracellular receptor
XX CC (iCR) and a transcription regulating domain (TRD). NBD is a polydactyl
XX CC zinc finger protein (ZFP), or a modular part of it, that interacts
XX CC specifically with a contiguous sequence of at least 3 nucleotides. The
XX CC fusion protein functions as a ligand-activated transcriptional regulator.
XX CC The fusion protein and the nucleic acid encoding it, are used to regulate
XX CC gene expression, particularly in gene therapy for treating malignant
XX CC cell proliferative diseases (e.g. colon cancer, prostate cancer,
XX CC renal-cell carcinoma) and non-malignant cell proliferative
XX CC diseases (e.g. psoriasis, pemphigus vulgaris, Behcet's syndrome and
XX CC lipid histiocytosis). The fusion protein and its DNA are also useful for
XX CC treating diseases caused by viruses in humans/plants, genetic and/or
XX CC acquired diseases. The fusion protein can be designed to target any
XX CC selected gene (endogenous or exogenous), and can be made to have
XX CC different selectivity or specificity for endogenous or exogenous ligands.
XX CC The present sequence is linker peptide present between zinc finger
XX CC domains of murine Cys2-His2 ZFP Zif268. The ZFP is used in the
XX CC fusion protein of the invention.
XX SQ Sequence 5 AA;

Query Match 71.4%; Score 5; DB 22; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGEKP 5
|||||
Db 1 TGEKP 5

RESULT 21
AAB62164
ID AAB62164 standard; peptide; 5 AA.

XX AC AAB62164;
XX DT 11-JUN-2001 (first entry)
XX DE Linker fragment for linking two zinc finger proteins.
XX KW Zinc finger protein; drug discovery; target validation; cancer; fungal;
XX KW functional genomic; genetic disease; protozoal; bacterial; infection;
XX KW viral; ischemia; vascular disease; arthritis; immunological disorder;
XX KW gene expression; linker.
XX OS Synthetic.
XX PN WO200119981-A2.
XX PD 22-MAR-2001.
XX PF 12-SEP-2000; 2000WO-US24897.
XX PR 14-SEP-1999; 99US-0395448.
XX PA (SANG-) SANGAMO BIOSCIENCES INC.
XX PI Case CC, Zhang L;
XX DR WPI; 2001-257881/26.
XX PT Identifying biological function of genes for functional genomics and
XX PT target validation applications, involves using recombinant zinc finger
XX PT proteins which specifically recognize genes with high efficacy -
XX PS Disclosure; Page 35; 86pp; English.
XX CC The invention provides methods of regulating gene expression using
XX CC recombinant zinc finger proteins. A method of identifying a biological
XX CC function of a candidate gene comprises selecting a candidate gene (F1),
XX CC providing a zinc finger protein (ZFP) that binds to a target site of F1,
XX CC culturing a cell under conditions, where ZFP contacts F1 and modulates
XX CC expression of F1, and assaying for a selected phenotype. The method is
XX CC used for identifying the biological function of a candidate gene
XX CC associated with a selected phenotype e.g. for drug discovery, target
XX CC validation and functional genomics. The functional genomic assays using
XX CC ZFPs allow for discovery of novel human and mammalian therapeutic
XX CC applications, including discovery of novel drugs for e.g. treatment of
XX CC genetic disease, cancer, fungal, protozoal, bacterial and viral
XX CC infection, ischemia, vascular disease, arthritis, or immunological
XX CC disorders. ZFPs and expression vectors encoding ZFPs can be administered
XX CC directly to a subject or cell for modulation of gene expression. Zinc
XX CC finger protein technology can be used to rapidly analyze differential
XX CC gene expression studies and to readily up or down regulate any candidate
XX CC target gene. Sequences AAB62163-170 represent linker fragments that are
XX CC used to link two zinc finger proteins.

Query Match 71.4%; Score 5; DB 22; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGEKP 5
|||||
Db 1 TGEKP 5

RESULT 22
AAE28170
ID AAE28170 standard; peptide; 5 AA.

XX AAE28170;
 XX 27-DEC-2002 (first entry)
 XX Linker peptide #2 used to link zinc finger proteins.
 DE
 XX Phenotype; zinc finger protein; ZFP; gene expression; tumour progression;
 KW transformation; inflammatory response; neurological disorder; linker.
 XX
 OS Synthetic.
 XX US2002094529-A1.
 PN
 XX 18-JUL-2002.
 PD
 XX 28-AUG-2001; 2001US-0941450.
 XX
 XX 14-SEP-1999; 99US-0395448.
 PR
 XX (CASE/) CASE C C.
 PA (URNO/) URNOV F.
 PA
 XX Case CC, Urnov F;
 PI
 XX WPI; 2002-673639/72.
 DR
 XX Identifying a gene, useful for exploring biological processes, e.g.
 PT transformation, tumor progression, or inflammatory response, comprises
 PT contacting a cell comprising a putative gene sequence with a zinc
 PT finger protein -
 PT
 XX Disclosure; Page 35; 46pp; English.
 PS
 XX The present invention relates to methods and compositions for identifying
 CC a gene. The method involves contacting a cell comprising a putative gene
 CC sequence (PGS) with an exogenous molecule that binds to and modulates the
 CC expression of the PGS, and assaying the cell for at least one selected
 CC phenotype, where if one or more of the selected phenotypes are observed,
 CC the PGS is identified as a gene. The methods include the use of exogenous
 CC molecules such as zinc finger proteins (ZFPs) which are capable of
 CC binding to and modulating expression of gene transcription, targeted to
 CC PGS, followed by assay of one or more selected phenotypes. The ZFPs are
 CC useful for determining the phenotypic consequences and function of gene
 CC expression. The method is useful to explore biological processes such as
 CC transformation, tumour progression, inflammatory response, neurological
 CC disorders, etc. The method can very rapidly identify genes that are up or
 CC down-regulated in different tissues or in response to specific stimuli.
 CC The present sequence is a linker peptide used to link two zinc finger
 CC proteins. This peptide is used to illustrate the method of the invention.
 XX
 XX Sequence 5 AA;
 SQ
 Query Match 71.4%; Score 5; DB 23; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TGEKP 5
 Db |||||
 1 TGEKP 5
 RESULT 23
 AAO22206
 ID AAO22206 standard; Peptide; 5 AA.
 XX
 AC AAO22206;
 XX
 XX 11-OCT-2002 (first entry)
 DT
 DE Zinc finger protein peptide linker SEQ ID No 5.
 DE
 XX Non-canonical zinc finger binding protein; ZFP; gene therapy.
 KW

XX Unidentified.
 OS
 XX WO200257293-A2.
 PN
 XX 25-JUL-2002.
 PD
 XX 22-JAN-2002; 2002WO-US01893.
 PF
 XX 22-JAN-2001; 2001US-263445P.
 PR
 PR 11-MAY-2001; 2001US-290716P.
 XX
 XX (SANG-) SANGAMO BIOSCIENCES INC.
 PA
 XX Rebar E, Jamieson A;
 PI
 XX WPI; 2002-566791/60.
 DR
 XX Non-canonical zinc finger binding protein for modulating gene
 PT expression comprises non-canonical zinc finger components that bind to
 PT a target sequence -
 PT
 XX Disclosure; Page 19; 63pp; English.
 PS
 XX The invention relates to an isolated, non-canonical (e.g., non-C2H2) zinc
 CC finger binding protein (ZFP) comprising one or more non-canonical zinc
 CC finger components that bind to a target sequence. A fusion polypeptide of
 CC the invention is useful for modulating expression of a gene. The non-
 CC canonical ZFP and its encoding polynucleotide, and a fusion protein
 CC comprising the non-canonical ZFP and its encoding polynucleotide can be
 CC used to treat disease. The non-canonical ZFP can be used in diagnostic
 CC assays and to link phenotype to expression of particular genes. The
 CC polynucleotide encoding the non-canonical ZFP can be used to treat
 CC disorders by gene therapy. This sequence represents a peptide relating to
 CC the zinc finger binding protein of the invention.
 XX
 XX Sequence 5 AA;
 SQ
 Query Match 71.4%; Score 5; DB 23; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TGEKP 5
 Db |||||
 1 TGEKP 5
 RESULT 24
 ABB83545
 ID ABB83545 standard; peptide; 5 AA.
 XX
 AC ABB83545;
 XX
 XX 27-SEP-2002 (first entry)
 DT
 XX Peptide linker #1.
 DE
 XX Zinc finger; stress tolerance; pathogen resistance;
 KW agrochemical; peptide linker.
 KW
 XX Unidentified.
 OS
 XX WO200257294-A2.
 PN
 XX 25-JUL-2002.
 PD
 XX 22-JAN-2002; 2002WO-US01906.
 PF
 XX 22-JAN-2001; 2001US-263445P.
 PR
 PR 11-MAY-2001; 2001US-290716P.
 XX
 XX (SANG-) SANGAMO BIOSCIENCES INC.
 PA
 XX

PI. Jamieson A, Li Q;
 XX WPI; 2002-566792/60.
 XX Modified plant zinc finger protein for modulating gene expression in a
 PT plant cell comprises zinc fingers that bind to a target site -
 XX
 XX Disclosure; Page 22; 50pp; English.
 XX
 XX The present invention relates to a modified plant zinc finger
 CC protein. This zinc finger protein is used to modulated gene
 CC expression in a plant cell. Nucleic acid encoding the zinc finger is
 CC expressed in plant cells to produce a plant with an altered phenotype
 CC relative to the wild-type plant. The altered phenotype is high in
 CC nutritional value, yield, stress tolerance, pathogen resistance,
 CC resistance to agrochemicals, production of pharmaceutical compounds or
 CC production of industrial chemicals. The present sequence is
 CC a peptide linker used to link the zinc finger to zinc finger
 CC proteins.
 XX
 XX
 SQ Sequence 5 AA;
 Query Match 71.4%; Score 5; DB 23; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TGEKP 5
 |||||
 Db 1 TGEKP 5
 |||||
 RESULT 25
 ABJ03944
 ID ABJ03944 standard; Peptide; 5 AA.
 AC ABJ03944;
 XX
 DT 25-SEP-2002 (first entry)
 DE Peptide linker SEQ ID NO: 211.
 XX
 KW Zinc finger protein; angiogenesis; vasculogenesis; ischaemia;
 KW diabetic retinopathy; psoriasis; arthropathy; cancer; tumour growth;
 KW gene therapy; antiatherosclerotic; vasotropic; antiarthritic; vulnerary;
 KW antulcer; cytostatic; antipsoriatic; antidiabetic; ophthalmological;
 KW osteopathic; antiinfertility.
 OS Synthetic.
 XX
 XX WO200246412-A2.
 XX
 PD 13-JUN-2002.
 XX
 PF 06-DEC-2001; 2001WO-US46861.
 XX
 PR 07-DEC-2000; 2000US-0733604.
 PR 12-DEC-2000; 2000US-0736083.
 PR 30-APR-2001; 2001US-0846033.
 XX
 PA (SANG-) SANGAMO BIOSCIENCES INC.
 XX
 PI Rebar E, Jamieson A, Liu Q, Liu P, Wolffe A, Eisenberg SP;
 PI Jarvis E;
 XX
 DR WPI; 2002-527918/56.
 XX
 PT New zinc finger protein that binds to target site in vascular
 PT endothelial growth factor gene, useful for modulating expression of the
 PT gene and for treating atherosclerosis, ischemia, arthritis, wound or
 PT ulcer -
 XX
 XX Disclosure; Page 187; 195pp; English.

CC The present invention relates to a zinc finger protein that binds to a
 CC target site in one or more vascular endothelial growth factor (VEGF)
 CC genes. The protein is useful for modulating expression of a VEGF gene,
 CC thereby regulating angiogenesis and vasculogenesis. This can be used to
 CC treat atherosclerosis, ischaemia, arthritis, wounds, ulcers, tumours,
 CC diabetic retinopathy or psoriasis. The present sequence is a peptide
 CC shown in the invention.
 XX
 SQ Sequence 5 AA;
 Query Match 71.4%; Score 5; DB 23; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TGEKP 5
 |||||
 Db 1 TGEKP 5
 |||||
 RESULT 26
 ABP48192
 ID ABP48192 standard; Peptide; 5 AA.
 XX
 AC ABP48192;
 XX
 DT 28-AUG-2002 (first entry)
 DE Peptide linker SEQ ID NO:2.
 XX
 KW Zinc finger protein; ZFP; DNA binding protein; zinc finger.
 XX
 OS Synthetic.
 XX
 FN WO200242459-A2.
 XX
 PD 30-MAY-2002.
 XX
 PF 20-NOV-2001; 2001WO-US43438.
 XX
 PR 20-NOV-2000; 2000US-0716637.
 XX
 PA (SANG-) SANGAMO BIOSCIENCES INC.
 XX
 LIU Q;
 XX
 DR WPI; 2002-500284/53.
 XX
 PT New zinc finger protein that binds to target site, useful in studying
 PT gene function and for human therapeutics and plant engineering.
 PT comprises first, second and third zinc fingers, ordered from N- to
 PT C-terminus -
 XX
 PS Disclosure; Page 15; 81pp; English.
 XX
 CC The present invention describes a zinc finger protein (I) that binds to
 CC a target site, comprising a first (F1), a second (F2), and a third (F3)
 CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
 CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
 CC and a third (S3) target sub-site. Also described are: (i) a polypeptide
 CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
 CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
 CC it binds to the S1 target sub-site, selecting the F2 zinc finger such
 CC that it binds to the S2 target sub-site, and selecting the F3 zinc
 CC finger such that it binds to the S3 target sub-site, thus designing (I)
 CC that binds to a target site. (I) is useful for recognition of triplet
 CC target sub-sites having the nucleotide G in the 5'-most position of the
 CC sub-site. (I) is useful in studying gene function, and for human
 CC therapeutics and plant engineering. (I), (II) or (III) is useful in
 CC therapeutic methods to modulate the expression of a target region within
 CC a subject, in diagnostic methods for sequence specific detection of
 CC target nucleic acid in a sample, and in assays to determined the
 CC phenotype and function of gene expression. (I) has improved affinity
 CC and specificity for their target sequences, as well as enhanced

CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
CC represent DNA target sequences and zinc finger peptides which are given
CC in the exemplification of the present invention.

XX
SQ Sequence 5 AA;
Query Match 71.4%; Score 5; DB 23; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5
| | | | |
Db 1 TGEKP 5

RESULT 27
ABP96165
ID ABP96165 standard; peptide; 5 AA.
XX
AC ABP96165;
XX
DT 15-MAY-2003 (first entry)
XX
DE Flexible peptide linker SEQ ID NO:30.
XX
KW Zinc finger binding domain; zinc finger nucleotide binding region;
KW nucleotide structure modulation; nucleotide function modulation;
KW gene therapy; linker.
XX
OS Synthetic.
XX
PN WO2003016496-A2.
XX
PD 27-FEB-2003.
XX
PF 20-AUG-2002; 2002WO-US26388.
XX
PR 20-AUG-2001; 2001US-313693P.
XX
PR 21-AUG-2001; 2001US-313864P.
XX
PA (SCRI) SCRIPPS RES INST.
XX
PI Barbas CF, Dreier B;
XX
DR WPI; 2003-268320/26.
XX
PT Novel zinc finger nucleotide binding polypeptide useful for regulating
PT nucleotide sequence expression of a prokaryotic, eukaryotic, or viral
PT gene -
XX
PS Claim 35; Page 18; 47pp; English.

CC The present invention describes an isolated and purified zinc finger
CC nucleotide binding polypeptide (I) comprising a nucleotide binding
CC region of from 5-10 amino acid residues, which region binds
CC preferentially to a target nucleotide of the formula CNN, where N is
CC A, G, C or T. Also described is a peptide (I) where the binding region
CC has an amino acid residue sequence with the same nucleotide binding
CC characteristics as any of the peptides in ABP96167 to ABP96191. The
CC present invention also describes: (1) a peptide composition (II)
CC comprising several of the above polypeptides, where the polypeptides are
CC operatively linked to each other; (2) an isolated and purified
CC polynucleotide (III) encoding (I); (3) an expression vector (IV)
CC containing (III); and (4) a host cell (V) transformed with (III) or
CC (I) can be used in gene therapy. (II) can be used for regulating the
CC expression of a nucleotide sequence that contains the sequence
CC 5'-(CNN)n-3', where n is 2-12, by exposing the nucleotide sequence to
CC (II), where the nucleotide sequence is a gene which is eukaryotic,
CC prokaryotic, or viral gene. The eukaryotic gene is a plant gene or a
CC mammalian gene, preferably a human gene, and the prokaryotic gene is a
CC bacterial gene. (II), (III) and (IV) are useful for modulating
CC nucleotide structure and/or function. The present sequence represents a
CC specifically claimed flexible peptide linker, which is given in the

CC exemplification of the present invention.

XX
SQ Sequence 5 AA;
Query Match 71.4%; Score 5; DB 24; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEKP 5
| | | | |
Db 1 TGEKP 5

RESULT 28
ABG74218
ID ABG74218 standard; peptide; 5 AA.
XX
AC ABG74218;
XX
DT 16-APR-2003 (first entry)
XX
DE Zinc finger protein fusion proteins, peptide linker #2.
XX
KW Zinc finger protein; vascular endothelial growth factor; VEGF; ZFP;
KW VP16 activation domain; VEGF3a/1; VEGF1; SV-40; NLS; cancer;
KW nuclear localization sequence; genetic disease; ischaemia; arthritis;
KW vascular disease; immunological disorder; infection; fruit ripening;
KW plant disease resistance; sugar composition; oil composition; HIV;
KW plant yield; colour; diabetic retinopathy; macular degeneration;
KW rheumatoid arthritis; psoriasis; human immunodeficiency virus infection;
KW sickle cell anaemia; Alzheimer's disease; muscular dystrophy;
KW neurodegenerative disease; vascular disease; cystic fibrosis; stroke;
KW plant nutritional characteristic; storage capability.
XX
OS Synthetic.
XX
PN US2002:60940-A1.
XX
PD 31-OCT-2002.
XX
PF 28-AUG-2001; 2001US-0942087.
XX
PR 12-JAN-1999; 99US-0229037.
XX
PA (CASE/) CASE C C.
PA (WOLF/) WOLFPE A.
PA (URNO/) URNOV F.
PA (LAI/) LAI A.
PA (SNOW/) SNOWDEN A.
PA (TANS/) TAN S.
PA (GREG/) GREGORY P.
XX
PI Case CC, Wolffe A, Urnov F, Lai A, Snowden A, Tan S, Gregory P;
XX
DR WPI; 2003-209226/20.
XX
PT Modulating the expression of an endogenous gene in a cell, e.g. for
PT treating genetic diseases and developing plants with altered
PT phenotypes, comprises contacting a target site in the gene with a zinc
PT finger protein -
XX
PS Disclosure; Page 18; 51pp; English.

CC The invention relates to modulating the expression of an endogenous
CC cellular gene in a cell, comprising contacting a target site in the gene
CC with a designed or selected zinc finger protein (ZFP), where the ZFP
CC comprises a functional domain, and thus modulating the expression of the
CC cellular gene. Effector plasmids expressing vascular endothelial growth
CC factors (VEGF) ZFPs fused to the VP16 activation domain were introduced
CC into cells and were shown to upregulate the VEGF gene. Eukaryotic
CC expression vectors were constructed that fuse the VEGF3a/1 and the VEGF1
CC ZFPs to the simian virus (SV)-40 nuclear localization sequence (NLS) and
CC VP16. The method is useful in treating genetic diseases, cancers,

CC ischaemia, vascular disease, arthritis, immunological disorders, and
 CC fungal, protozoal, bacterial and viral infections. The method is also
 CC useful in functional genomics assays, and for developing plants with
 CC altered phenotypes, including disease resistance, fruit ripening, sugar
 CC and oil composition, yield and colour. The method is also useful for
 CC treating diabetic retinopathy, macular degeneration, rheumatoid
 CC arthritis, psoriasis, human immunodeficiency virus (HIV) infection,
 CC sickle cell anaemia, Alzheimer's disease, muscular dystrophy,
 CC neurodegenerative diseases, vascular disease, cystic fibrosis and stroke.
 CC ZFP proteins are administered to engineer plants for traits including
 CC increased disease resistance, modification of structural and storage
 CC polysaccharides, flavours, proteins and fatty acids, fruit ripening,
 CC yield, colour, nutritional characteristics and improved storage
 CC capability. The present sequence is a peptide linker suitable for
 CC use in linking the ZFP to a protein or domain of interest.

SQ Sequence 5 AA;
 Query Match 71.4%; Score 5; DB 24; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGEKP 5
 |||||
 Db 1 TGEKP 5

RESULT 29
 AAE30436
 ID AAE30436 standard; peptide; 5 AA.

AC AAE30436;
 DT 24-FEB-2003 (first entry)

DE Peptide linker #1 used to link zinc finger proteins.

KW DNA binding specificity; zinc finger protein; specificity optimisation;
 ZFP; gene expression.

OS Unidentified.

PN WO200277227-A2.

PP 03-OCT-2002.

PF 20-NOV-2001; 2001WO-US43568.

PR 20-NOV-2000; 2000US-0716637.

PA (SANG-) SANGAMO BIOSCIENCES INC.

PI Eisenberg SP, Liu Q, Jamieson A, Rebar E;

DR WPI; 2003-029936/02.

PT Enhancing the binding specificity of a zinc finger protein, comprises
 substituting amino acids in the protein for residues in the target
 sequence to make a modified binding protein -

PS Disclosure; Page 19; 55pp; English.

CC The present invention relates to a novel method of enhancing the binding
 CC specificity of a binding protein (zinc finger protein (ZFP)). The method
 CC involves substituting one or more amino acids at positions in the binding
 CC protein that affect the specificity of the binding protein for residues
 CC in the target sequence to make a modified binding protein. The method is
 CC useful in optimising the specificity of a binding protein, in modulating
 CC the expression of a target gene in a subject and in diagnostic methods
 CC for sequence-specific detection of a target nucleic acid in a sample.
 CC The present sequence is a peptide linker used to link zinc finger
 CC proteins.

XX

SQ Sequence 5 AA;
 Query Match 71.4%; Score 5; DB 24; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGEKP 5
 |||||
 Db 1 TGEKP 5

RESULT 30
 ABG73580

ID ABG73580 standard; Peptide; 6 AA.

AC ABG73580;

DT 27-FEB-2003 (first entry)

DE Human zinc finger protein 11.55-associated peptide #2.

KW Human; zinc finger protein 11.55; malignant tumour; haemopathy;
 human immunodeficiency virus; HIV; infection; immunological disease;
 inflammation.

OS Homo sapiens.

PH Key Location/Qualifiers

FT Misc-difference 6 /label= Tyr, Phe

PN CN1352067-A.

PD 05-JUN-2002.

PF 02-NOV-2000; 2000CN-0127161.

PR 02-NOV-2000; 2000CN-0127161.

PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.

PI Mao Y, Xie Y;

DR WPI; 2002-658695/71.

PT New human zinc finger protein 11.55 polypeptide for treating malignant
 tumors, hemopathy, human immunodeficiency virus infection,
 immunological diseases and various inflammations -

PS Disclosure; Page 2 (Disclosure); 33pp; Chinese.

CC This invention describes the novel human zinc finger protein 11.55,
 CC polynucleotides encoding this polypeptide and a DNA recombination process
 CC to produce the polypeptide. The invention also discloses applying the
 CC polypeptide in treating various diseases, such as malignant tumours,
 CC haemopathy, human immunodeficiency virus (HIV) infection, immunological
 CC diseases and various inflammations. This sequence represents a peptide
 CC associated with the human zinc finger protein 11.55 described in the
 CC method of the invention.

SQ Sequence 6 AA;

Query Match 71.4%; Score 5; DB 23; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGEKP 5
 |||||
 Db 1 TGEKP 5

RESULT 31
 AAY98576

ID AAY98576 standard; Peptide; 9 AA.
 XX AC AAY98576;
 XX DT 31-JUL-2000 (first entry)
 XX WT1 derived immunogenic peptide SEQ ID NO:91.
 DE WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
 XX KW metastatic disease; mouse; human; Wilm's tumour; immune response;
 KW vaccine.
 XX OS Homo sapiens.
 XX PN WO200018795-A2.
 XX PD 06-APR-2000.
 XX PF 30-SEP-1999; 99WO-US22819.
 XX PR 30-SEP-1998; 98US-0164223.
 XX PR 25-MAR-1999; 99US-0276484.
 XX PA (CORI-) CORIXA CORP.
 XX PA (GAIG/) GAIGER A.
 XX PI Gaiger A, Cheever M;
 XX DR WPI; 2000-293107/25.
 XX Novel polypeptides comprising an immunogenic portion of a native WT1
 PT polypeptide, useful for inhibiting the development of malignant
 PT diseases associated with WT1 expression e.g. leukemia or cancer -
 XX Claim 4; Page 157; 193pp; English.
 XX The present invention describes polypeptides (I) comprising an
 CC immunogenic portion of a native Wilm's Tumour gene product polypeptide,
 CC WT1, (or variants of the immunogenic portion retaining the ability to
 CC react with WT1-specific antisera and/or T-cell lines or clones) and
 CC comprising 16 consecutive amino acids (aa) or less of a native WT1
 CC polypeptide. The polypeptides are useful therapeutically and to
 CC manufacture medicaments for enhancing/inducing an immune response in
 CC patients. The polypeptides, mimetics or polynucleotides can be included
 CC with a carrier/exipient in pharmaceutical compositions or with a
 CC non-specific immune response enhancer (e.g. an adjuvant or enhancer
 CC preferentially enhancing a T cell response) in vaccines. Pharmaceutical
 CC compositions and vaccines can be administered to human patients to
 CC enhance or induce an immune response specific for WT1 or a cell
 CC expressing WT1, useful to inhibit the development of malignant diseases
 CC associated with WT1 expression, e.g. leukemia (especially acute/chronic
 CC myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
 CC breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
 CC AAU98501 to AAU98811 represent polypeptide sequences, and AAU13848 to
 CC AAU13862 represent PCR primers, used in the exemplification of the
 CC present invention.
 XX Sequence 9 AA;
 SQ Query Match 71.4%; Score 5; DB 21; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GEKPY 6
 Db |||||
 1 GEKPY 5
 RESULT 32
 AAU68675
 ID AAU68675 standard; Peptide; 9 AA.
 XX AAU68675;
 AC

XX 16-JAN-2002 (first entry)
 XX Human Wilm's tumour protein, WT1, antigenic peptide #70.
 DE Human; Wilm's tumour; WT1; pleural mesothelioma; antigen;
 KW leukaemia; acute myeloid leukaemia; AML; chronic myeloid leukaemia; CML;
 KW acute lymphocytic leukaemia; ALL; myelodysplastic syndromes;
 XX myeloproliferative syndrome; cancer; cytostatic.
 OS Homo sapiens.
 XX PN WO200162920-A2.
 XX PD 30-AUG-2001.
 XX PF 22-FEB-2001; 2001WO-US05702.
 XX PR 22-FEB-2000; 2000US-184070P.
 XX PA (CORI-) CORIXA CORP.
 XX PI Cheever MA, Gaiger A;
 XX DR WPI; 2001-648218/74.
 XX Composition for the treatment of mesothelioma comprises specific
 PT peptides i.e. Wilm's tumour antigen polypeptide derived antigenic
 PT fragments -
 XX Claim 1; Page 217; 242pp; English.
 XX The invention relates to the use of a composition comprising at least a
 CC first isolated peptide, of between 9 and 40 amino acids or a first
 CC nucleic acid, encoding the peptide, in the manufacture of a medicament
 CC for treating or preventing mesothelioma. The peptides are antigenic
 CC peptides derived from the Wilm's tumour protein WT1. The composition is
 CC useful for the treatment of mesothelioma, Wilm's tumour, preferably
 CC pleural mesothelioma and other WT1 associated malignancies e.g.
 CC leukaemia (including acute myeloid leukaemia, ALL, chronic myeloid
 CC leukaemia, CML, acute lymphocytic leukaemia, ALL, and childhood ALL),
 CC myelodysplastic syndromes, myeloproliferative syndromes and cancers
 CC (e.g. breast, testicular, prostate, lung and ovarian) in mammals,
 CC preferably humans. The present sequence is an antigenic peptide of
 CC the invention derived from human WT1.
 XX Sequence 9 AA;
 SQ Query Match 71.4%; Score 5; DB 22; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GEKPY 6
 Db |||||
 1 GEKPY 5
 RESULT 33
 AAG61908
 ID AAG61908 standard; Peptide; 9 AA.
 XX AAG61908;
 AC AAG61908;
 XX 06-JUL-2001 (first entry)
 DT Human WT1 immunogenic peptide SEQ ID NO: 91.
 DE Human; mouse; immunotherapy; cancer; leukaemia; WT1; Wilm's tumour gene;
 XX KW chromosome 11p13; zinc finger transcription factor.
 XX OS Homo sapiens.
 XX PN WO200125273-A2.

XX' 12-APR-2001.
 XX PD
 XX PF 04-OCT-2000; 2000WO-US27465.
 XX PR 04-OCT-1999; 99US-0157459.
 XX PA (CORI-) CORIXA CORP.
 XX PI Skeiky YAM, Xu J, Cheever MA, Reed SG;
 XX DR WPI; 2001-328324/34.
 XX PT Polypeptide comprising part of the Wilm's Tumour gene product sequence is
 XX used in the diagnosis and treatment of malignant diseases e.g. leukemia
 XX PT and cancer associated with WT1 -
 XX PS Claim 4; Page 169; 228pp; English.
 XX CC The present invention describes compositions comprising peptides derived
 XX from the Wilm's tumour protein WT1 and methods for their use in treating
 XX malignant diseases. Peptides derived from both the murine and human WT1
 XX proteins are provided. The human WT1 gene is found on chromosome 11p13,
 XX and the protein was shown to be a zinc finger transcription factor. The
 XX immunogenic peptides of the invention are particularly useful in the
 XX diagnosis and treatment of cancer and leukaemia. The present sequence is
 XX a polypeptide described in the exemplification of the invention.
 XX SQ Sequence 9 AA;
 Query Match 71.4%; Score 5; DB 22; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 GEXPY 6
 Db 1 GEXPY 5
 RESULT 34
 ABG33145
 ID ABG33145 standard; Peptide; 9 AA.
 XX AC
 XX DT 15-JUL-2002 (first entry)
 XX DE Human WT1 immunogenic peptide #68.
 XX KW Human; mouse; cytostatic; immunostimulant; WT1; cancer;
 XX KW immune response.
 XX OS Homo sapiens.
 XX PN WO200228414-A1.
 XX PD 11-APR-2002.
 XX PF 03-OCT-2001; 2001WO-US31139.
 XX PR 06-OCT-2000; 2000US-0684361.
 XX PR 09-OCT-2000; 2000US-0685830.
 XX PR 15-FEB-2001; 2001US-0785019.
 XX PR 24-AUG-2001; 2001US-0938864.
 XX PA (CORI-) CORIXA CORP.
 XX PA (GAIG/) GAIGER A.
 XX PI Gaiger A, McNeill PD, Smithgall M, Moulton G, Vedvick TS;
 XX PI Sleath PR, Mosseall S, Evans L, Spies AG, Boydston J;
 XX DR WPI; 2002-352217/38.

PT Novel isolated WT1 polynucleotide, and encoded polypeptide, useful for
 PT treating and diagnosing cancer in a patient -
 XX PS Example 4; Page 179; 260pp; English.
 XX CC The invention relates to an isolated WT1 polynucleotide (I) and
 XX CC polypeptide encoded by (I). The WT1 polynucleotides and polypeptides
 XX CC are used for treating and detecting cancer in a patient, and for
 XX CC stimulating an immune response in patient. ABG33070-ABG33405
 XX CC represent WT1 amino acid sequences of the invention.
 XX SQ Sequence 9 AA;
 Query Match 71.4%; Score 5; DB 23; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 GEXPY 6
 Db 1 GEXPY 5
 RESULT 35
 AAR99267
 ID AAR99267 standard; Peptide; 10 AA.
 XX AC
 XX AC AAR99267;
 XX DT 13-DEC-1996 (first entry)
 XX DE Leukocyte cell-derived chemotaxin LECT2a peptide 2.
 XX KW LECT2a; LECT2b; leukocyte cell-derived chemotaxin; cytokine;
 XX KW neutrophil; chemotaxis; cancer; therapy; diagnosis.
 XX OS Homo sapiens.
 XX PN EP723016-A2.
 XX PD 24-JUL-1996.
 XX PF 27-NOV-1995; 95EP-0402670.
 XX PR 28-NOV-1994; 94JP-0293233.
 XX PA (SUZU/) SUZUKI K.
 XX PI Amiga S, Mizuno S, Suzuki K, Yamakawa Y, Yamagoe S;
 XX DR WPI; 1996-335477/34.
 XX PT New human leukocyte cell-derived chemotaxin cpds. - used for the
 XX PT detection and therapy of cancer and disorders of the cytokine
 XX PT network
 XX PS Claim 1; Page 11; 17pp; English.
 XX CC LECT2a peptides 1-5 (AAR99266-70) comprise tryptic peptides of novel
 XX CC human leukocyte cell-derived chemotaxin LECT2a, in order from the
 XX CC N-terminal sequence. They were obtd. by tryptic digestion of 15
 XX CC kDa LECT2a isolated from the phytohaemagglutinin-activated T-cell
 XX CC leukaemia cell line SKW-3. LECT2a and LECT2b (AAR99271) have
 XX CC neutrophil chemotactic activity. They are useful for the diagnosis,
 XX CC therapy and prediction of cancer and disorders of the cytokine
 XX CC network.
 XX SQ Sequence 10 AA;
 Query Match 71.4%; Score 5; DB 17; Length 10;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 3 EKPYP 7

```

Db      |||||
        5 EKPXK 9

RESULT 36
AAW78383
ID AAW78383 standard; peptide; 4 AA.
XX AC
XX AAW78383;
XX 11-MAY-1999 (first entry)
DT XX
XX Synthetic Cys2-his2 zinc finger linker peptide #1.
DE XX
XX Zinc finger; target sequence; binding assay; mutant;
KW phosphorylation site; functional domain.
XX AC
XX Synthetic.
XX WO9853060-A1.
XX 26-NOV-1998.
XX 26-MAY-1998; 98WO-GB01516.
XX 23-MAY-1997; 97GB-0010809.
XX (MEDI-) MEDICAL RES COUNCIL.
XX Choo Y, Isalan M, Klug A;
XX WPI; 1999-045309/04.
XX Rules for designing zinc finger nucleic acid binding proteins
PT specific for any base quadruplet - relate bases in the quadruplet to
PT specific amino acids in the alpha-helical binding motif, used to
PT detect target nucleic acids, e.g. for identification of mutants and
PT phosphorylation sites
XX Claim 9; Page 41; 57pp; English.
XX This sequence represent a linker peptide for a generic zinc finger of
CC the Cys2-His2 zinc finger (2P) class (AAW78382). The 2P are generated
CC so that they able to bind a nucleic acid quadruplet in a target
CC sequence, where binding to base 4 of the quadruplet by an alpha-helical
CC ZF binding motif is determined as: (a) if base 4 is A, then position +6
CC in the helix is Gln and position +2 is not Asp (+2 indicates a residue
CC present in an adjacent, C-terminal ZF) and (b) if base 4 is C, then
CC position +6 may be any residue provided +2 is not Asp. The 2P are used
CC to detect target nucleic acids in a binding assay, e.g. for
CC identification of mutants (they can differentiate between single bp
CC changes in the target) or potential phosphorylation sites, and to
CC characterise functional domains of a protein.
XX
XX SQ Sequence 4 AA;
Query Match 57.1%; Score 4; DB 20; Length 4;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEK 4
Db |||||
1 TGEK 4

RESULT 37
AAW84300
ID AAW84300 standard; Peptide; 4 AA.
XX AC
XX AAW84300;
XX 18-MAR-1999 (first entry)
DT XX
XX Zinc finger protein binding motif linker #1.
KW TFIIIA/Zif-VP16; sub-cellular trafficking; signal transduction;
KW diagnostic; gene therapy; agricultural technology; zinc finger.
XX OS
XX Synthetic.
XX WO200153479-A2.
XX 26-JUL-2001.
XX

```

```

DE Linker used in the zinc fingers of the invention.
XX Zinc finger; nucleic acid binding protein; NABP; linker;
KW Cys2-His2 zinc finger; detection; gene therapy; gene delivery.
XX OS
XX Synthetic.
XX WO9853059-A1.
XX 26-NOV-1998.
XX 26-MAY-1998; 98WO-GB01514.
XX 23-MAY-1997; 97GB-0010807.
XX (MEDI-) MEDICAL RES COUNCIL.
XX Choo Y, Isalan M, Klug A;
XX WPI; 1999-045308/04.
XX Preparation of nucleic acid binding proteins - by designing protein
PT sequences of a Cys2-His2 zinc finger class based on a nucleic acid
PT base triplet in a target nucleic acid sequence
XX Claim 9; Page 44; 62pp; English.
XX The present sequence represents a linker used in the zinc fingers
CC of the invention. The specification describes a method for preparing
CC a nucleic acid binding protein (NABP) of the Cys2-His2 zinc finger
CC class capable of binding to a nucleic acid base triplet in a target
CC nucleic acid sequence. Binding to the 5' base of the triplet by an
CC alpha-helical zinc finger nucleic acid binding motif in the protein
CC is determined as follows: (a) if the 5' base in the triplet is A,
CC then position +6 in the alpha-helix is Glu, Asn or Val; (b) if the
CC 5' base in the triplet is C, then position +6 in the alpha-helix is
CC Ser, Thr, Val, Ala, Glu or Asn. The methods can be used for designing
CC a protein which is capable of binding to any predefined nucleic acid
CC sequence. The NABPs can be used for the detection of target nucleic
CC acid molecules. They can also be used in gene therapy, e.g. for
CC the delivery of functional genes into defective genes, or the delivery
CC of nonsense nucleic acid to disrupt undesired nucleic acid.
XX
XX SQ Sequence 4 AA;
Query Match 57.1%; Score 4; DB 20; Length 4;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEK 4
Db |||||
1 TGEK 4

RESULT 38
AAU05194
ID AAU05194 standard; Protein; 4 AA.
XX AC
XX AAU05194;
XX 24-OCT-2001 (first entry)
DT XX
XX Zinc finger protein binding motif linker #1.
DE XX
KW TFIIIA/Zif-VP16; sub-cellular trafficking; signal transduction;
KW diagnostic; gene therapy; agricultural technology; zinc finger.
XX OS
XX Synthetic.
XX WO200153479-A2.
XX 26-JUL-2001.
XX

```


PF 18-JAN-2001; 2001WO-GB00187.
 XX
 PR 24-JAN-2000; 2000GB-0001578.
 PR 24-JAN-2000; 2000GB-0001582.
 PR 30-MAY-2000; 2000WO-GB02071.
 PR 30-MAY-2000; 2000WO-GB02080.
 PR 07-DEC-2000; 2000GB-0029901.
 XX
 PA (GEND-) GENDAQ LTD.
 XX
 PI Choo Y, Ullman CG, Moore M;
 XX
 DR WPI; 2001-457607/49.

PT Selecting switching system components comprising first molecule, second
 PT molecule and ligand, involves determining degree of binding between the
 PT molecules in the presence and absence of the ligand -
 XX

PS Disclosure; Page 51; 193pp; English.

CC The invention relates to a method of selecting components of a switching
 CC system comprising first molecule (M1), selecting (I) a ligand (L2) which
 CC is capable of modulating the interaction between a nucleic acid binding
 CC molecule (NBM) and a target nucleic acid (TN). NBM, TN and the ligands
 CC are useful for regulating a biological process selected from sub-cellular
 CC trafficking and signal transduction, and for regulating transcription or
 CC translation from a nucleic acid sequence comprising TN to which NBM binds
 CC in a manner modulatable by the ligand. They are also useful for
 CC modulating a biological process affecting one or more genes, such as in
 CC transcription, translation, phosphorylation, replication, transport,
 CC splicing, integration and recombination, in a host cell. The polypeptides
 CC or ligands selected from (S) are useful for regulating a biological
 CC process involving binding of first polypeptide with second polypeptide,
 CC selected from sub-cellular trafficking and signal transduction. (I) is
 CC also useful for screening any or all the components of the gene switch
 CC system or protein switch system, and for modulating the interaction
 CC between the components of the switching system. The binding molecules are
 CC useful as diagnostic and research tools, and for switching or modulating
 CC gene expression, especially in gene therapy applications and agricultural
 CC technology applications. The ligands are also useful in the treatment of
 CC diseases. Multiple target nucleic acids or polypeptides could be
 CC used in a single selective step, thus enabling multiple nucleic acid or
 CC polypeptide binding molecules to be isolated simultaneously, even in
 CC the same physical vessel. The present sequence represents the zinc finger
 CC protein binding motif linker #1 used in the method of the invention.

SQ Sequence 4 AA;

Query Match 57.1%; Score 4; DB 22; Length 4;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEK 4
 ||||
 Db 1 TGEK 4

RESULT 39
 AAB85405
 ID AAB85405 standard; peptide; 4 AA.

XX AAB85405;

DT 17-SEP-2001 (first entry)

DE Peptide linker for generating zinc finger fusion constructs.

XX Nucleic acid binding polypeptide; repressor domain; cardiant; nootropic;
 KW circulatory active; anti-inflammatory; dermatological; neuroprotective;
 KW cerebroprotective; antibacterial; antifungal; antiviral; antirheumatic;
 KW osteopathic; gene therapy; zinc finger; binding site; linker.
 XX
 OS Synthetic.

XX WO200153480-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 19-JAN-2001; 2001WO-GB00202.
 XX
 PR 24-JAN-2000; 2000GB-0001582.
 PR 30-MAY-2000; 2000GB-0013102.
 PR 30-MAY-2000; 2000GB-0013103.
 PR 30-MAY-2000; 2000GB-0013104.
 XX
 PA (GEND-) GENDAQ LTD.

XX Choo Y, Klug A, Moore M;
 XX
 DR WPI; 2001-451906/48.

PT Nucleic acid binding polypeptide, used to identify nucleic acids and
 PT treat inflammatory, neurological, and dermatological disease, comprises
 PT a repressor domain and several nucleic acid binding domains linked by
 PT non-canonical linker(s) -
 PS Claim 23; Page 104; 142pp; English.

CC The invention relates to a nucleic acid (NA) binding polypeptide (I)
 CC comprising a repressor domain and several NA binding domains (BDs) linked
 CC by at least one non-canonical linker. (I) may be used to identify NAs in
 CC a complex mixture, to differentiate single base pair changes in NAs, in
 CC the manufacture of chimeric restriction enzymes, to produce knock out
 CC organisms, and in the treatment of diseases such as: cardiovascular,
 CC inflammatory, metabolic, infectious, neurological, rheumatological,
 CC genetic, dermatological, and musculoskeletal diseases. The invented
 CC methods are used to produce novel NA binding polynucleotides and to
 CC modify existing NA binding polynucleotides comprising several NA BDs.
 CC The novel polypeptide comprises several nucleic acid binding domains
 CC linked by linker sequences. The invented polypeptide is therefore able
 CC to span longer or variable gaps, and a greater number of gaps, between
 CC DNA binding subsites. Sequences AAB85405-414 represent peptide linkers
 CC used in the construction of zinc finger fusion constructs.

SQ Sequence 4 AA;

Query Match 57.1%; Score 4; DB 22; Length 4;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GEXP 5
 ||||
 Db 1 GEXP 4

RESULT 40
 AABG79360
 ID AABG79360 standard; Peptide; 4 AA.

XX AABG79360;

DT 15-NOV-2002 (first entry)

DE Canonical linker peptide #2.

XX Receptor; virus infection; immune response; autoimmune disease;
 KW TNFR1 receptor; inflammation; autoimmune encephalomyelitis;
 KW rheumatoid arthritis; myocarditis; canonical linker.
 XX
 OS Synthetic.

XX WO200257308-A2.

XX 25-JUL-2002.

PF 22-JAN-2002; 2002WO-GB00246.

XX PR 22-JAN-2001; 2001GB-0001576.
 XX PA 07-FEB-2001; 2001GB-0003032.
 XX (SANG-) SANGAMO BIOSCIENCES INC.
 XX Moore M, Isalan M, Reynolds L, Ullman C, Girdlestone J;
 XX Demaison C, Choo Y;
 XX WPI; 2002-590720/63.
 XX New polypeptide capable of binding nucleic acids for treating or
 PT preventing a disease caused by a virus or a disease associated with an
 PT immune response, particularly autoimmune diseases e.g. inflammation or
 PT rheumatoid arthritis -
 XX Disclosure; Page 35; 105pp; English.
 XX The present invention relates to a new polypeptide capable of binding to
 CC a nucleic acid comprising a receptor nucleotide sequence which can
 CC function as a receptor for virus infection or is involved in an immune
 CC response. The molecules of the invention are useful in the treatment or
 CC prevention of a disease caused by a virus or of a disease associated with
 CC an immune response, particularly an autoimmune disease. The zinc finger
 CC polypeptide or a nucleic acid encoding such a polypeptide can be used to
 CC modulate transcription of a receptor nucleotide sequence. The nucleic
 CC acid polypeptide capable of binding to a nucleic acid sequence comprising
 CC a receptor nucleotide sequence, where the receptor is capable of
 CC functioning as a receptor for infection by the virus or is involved in
 CC an immune response, or a nucleic acid encoding such a polypeptide, can
 CC be used in the preparation of a medicament for use in the treatment or
 CC prevention of a disease caused by a virus or of a disease associated
 CC with an immune response. The polypeptides may further be used to treat or
 CC prevent various diseases or syndromes associated with or caused by
 CC malfunction of the receptor as the TNFR1 receptor, such as autoimmune
 CC diseases including inflammation, autoimmune encephalomyelitis, rheumatoid
 CC arthritis or myocarditis. The present amino acid sequence represents a
 CC canonical linker peptide sequence, as described in the invention.
 XX Sequence 4 AA;
 SQ Query Match 57.1%; Score 4; DB 23; Length 4;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GEKP 5
 Db 1 GEKP 4
 RESULT 41
 ID ABP57927 standard; peptide; 4 AA.
 XX AC ABP57927;
 XX 21-FEB-2003 (first entry)
 XX Zinc finger protein flexible linker peptide #2.
 XX Zinc finger; transgenic; transgene; cardiant; cytostatic; vulnery;
 KW nontropic; neuroprotective; osteopathic; antiarthritis; gene therapy;
 KW protein therapy; hypertrophic cardiomyopathy; cancer; emphysema; burn;
 KW Alzheimer's disease; osteoarthritis; linker peptide.
 XX OS Synthetic.
 XX WO200279419-A2.
 XX 10-OCT-2002.
 XX 28-MAR-2002; 2002WO-US09703.

PR 28-MAR-2001; 2001GB-0007757.
 XX (SANG-) SANGAMO BIOSCIENCES INC.
 XX Choo Y, Girdlestone J, Moore M;
 XX WPI; 2003-103268/09.
 XX New transgenic animal for studying genetics and diseases in animal
 PT models, comprises a nucleic acid binding polypeptide which binds and
 PT modulates the expression of a target gene, and a nucleotide sequence
 PT encoding the polypeptide -
 XX Disclosure; Page 38; 90pp; English.
 XX The invention relates to a novel transgenic non-human animal comprising a
 CC heterologous nucleic acid binding polypeptide which binds to a target
 CC gene and modulates its expression, where the nucleic acid binding
 CC polypeptide is encoded by a transgene, and where the expression of a
 CC target gene in at least one cell is modulated compared to a
 CC non-transgenic animal. The transgene of the invention has cardiant,
 CC cytostatic, vulnery, nontropic, neuroprotective, osteopathic, and
 CC antiarthritis activity. The polynucleotide may have a use in gene
 CC therapy, and protein therapy. The transgenic animal is useful in studying
 CC the relationship between genetics and disease in animal models, and the
 CC effect of therapeutic treatments for these diseases. The transgenic
 CC animal models are also used for studying developmental processes. The
 CC polypeptides and the polynucleotides may be used in treating diseases,
 CC such as hypertrophic cardiomyopathy, cancer, emphysema, burns,
 CC Alzheimer's disease and osteoarthritis. The present sequence represents a
 CC flexible linker peptide, used to join the last residue of the alpha-helix
 CC in a zinc finger and the first residue of the beta-sheet in the next zinc
 CC finger.
 XX Sequence 4 AA;
 SQ Query Match 57.1%; Score 4; DB 24; Length 4;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GEKP 5
 Db 1 GEKP 4
 RESULT 42
 ID ABP55221 standard; Peptide; 4 AA.
 XX AC ABP55221;
 XX 28-JAN-2003 (first entry)
 XX Linker peptide sequence #3.
 XX Gene regulation; expression; nucleic acid binding protein; cytostatic;
 KW nephrotropic; gene therapy; kidney failure; cancer.
 XX OS Synthetic.
 XX WO200274996-A1.
 XX 26-SEP-2002.
 XX 19-MAR-2002; 2002WO-US08554.
 XX 19-MAR-2001; 2001GB-0006786.
 XX (SANG-) SANGAMO BIOSCIENCES INC.
 XX Girdlestone J, England N, Demaison C;
 XX WPI; 2003-058340/05.

XX Regulating expression of a nucleic acid sequence in a primary cell for
 PT treating or preventing a disease e.g., cancer, comprises contacting the
 PT nucleic acid binding polypeptide with the nucleic acid sequence -
 XX
 PS Disclosure; Page 29; 81pp; English.
 XX
 CC The present invention describes a method (M1) for regulating expression
 CC of a nucleic acid sequence in a primary cell comprising providing a
 CC nucleic acid binding polypeptide capable of binding to the nucleic acid
 CC sequence and contacting the nucleic acid binding polypeptide with the
 CC nucleic acid sequence in the primary cell. Also described: (1) a nucleic
 CC acid binding polypeptide (I) capable of binding to, and regulating the
 CC expression of, a nucleic acid sequence in a primary cell; (2) a primary
 CC cell (II) comprising an exogenous nucleic acid binding polypeptide;
 CC (3) a pharmaceutical composition (III) comprising the polypeptide or the
 CC primary cell and a carrier or diluent; (4) treating (M2) or preventing a
 CC disease; or (5) expressing (M3) an exogenous nucleic acid binding
 CC polypeptide in a primary cell. (I) has cytostatic and nephrotoxic
 CC activities, and can be used in gene therapy. The method is useful for
 CC treating or preventing a disease e.g., kidney failure or cancer. The
 CC present sequence represents a linker peptide which is used in the
 CC exemplification of the present invention.

XX Sequence 4 AA;
 SQ
 Query Match 57.1%; Score 4; DB 24; Length 4;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GEXP 5
 Db 1 GEXP 4

RESULT 43
 AAB85409
 ID AAB85409 standard; peptide; 5 AA.
 XX
 AC AAB85409;
 XX
 DT 17-SEP-2001 (first entry)
 XX
 DE Peptide linker for generating zinc finger fusion constructs.

XX Nucleic acid binding polypeptide; repressor domain; cardiant; nootropic;
 KW circulatory active; anti-inflammatory; dermatological; neuroprotective;
 KW cerebroprotective; antibacterial; antifungal; antiviral; antirheumatic;
 KW osteopathic; gene therapy; zinc finger; binding site; linker.

XX Synthetic.
 OS
 XX WO200153480-A1.
 PN
 XX 26-JUL-2001.
 PD
 XX 19-JAN-2001; 2001WO-GB00202.
 PF
 XX 24-JAN-2000; 2000GB-0001582.
 PR 30-MAY-2000; 2000GB-0013102.
 PR 30-MAY-2000; 2000GB-0013103.
 PR 30-MAY-2000; 2000GB-0013104.

XX (GEND-) GENDAO LTD.

XX Choo Y, Klug A, Moore M;
 XX WPI; 2001-451906/48.
 DR

XX Nucleic acid binding polypeptide, used to identify nucleic acids and
 PT treat inflammatory, neurological, and dermatological disease, comprises
 PT a repressor domain and several nucleic acid binding domains linked by
 PT non-canonical linker(s) -

XX Claim 24; Page 105; 142pp; English.
 PS
 XX The invention relates to a nucleic acid (NA) binding polypeptide (I)
 CC comprising a repressor domain and several NA binding domains (BDs) linked
 CC by at least one non-canonical linker. (I) may be used to identify NAs in
 CC a complex mixture, to differentiate single base pair changes in NAs, in
 CC the manufacture of chimeric restriction enzymes, to produce knock out
 CC organisms, and in the treatment of diseases such as: cardiovascular,
 CC inflammatory, metabolic, infectious, neurological, rheumatological,
 CC genetic, dermatological, and musculoskeletal diseases. The invented
 CC methods are used to produce novel NA binding polynucleotides and to
 CC modify existing NA binding polynucleotides comprising several NA BDs.
 CC The novel polypeptide comprises several nucleic acid binding domains
 CC linked by linker sequences. The invented polypeptide is therefore able
 CC to span longer or variable gaps, and a greater number of gaps, between
 CC DNA binding subunits. Sequences AAB85405-414 represent peptide linkers
 CC used in the construction of zinc finger fusion constructs.

XX Sequence 5 AA;

XX Query Match 57.1%; Score 4; DB 22; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GEXP 5
 Db 2 GEXP 5

RESULT 44
 AABG79363
 ID AABG79363 standard; Peptide; 5 AA.

XX AABG79363;
 XX
 DT 15-NOV-2002 (first entry)
 XX
 DE Canonical linker peptide #5.

XX Receptor; virus infection; immune response; autoimmune disease;
 KW TNFRI receptor; inflammation; autoimmune encephalomyelitis;
 KW rheumatoid arthritis; myocarditis; canonical linker.

XX Synthetic.
 OS
 XX WO200257308-A2.

PN 25-JUL-2002.

PD 22-JAN-2002; 2002WO-GB00246.

PF 22-JAN-2001; 2001GB-0001576.

PR 07-FEB-2001; 2001GB-0003032.

XX (SANG-) SANGAMO BIOSCIENCES INC.

XX Moore M, Isalan M, Reynolds L, Ullman C, Girdlestone J;
 PI Demaison C, Choo Y;

XX WPI; 2002-590720/63.

XX New polypeptide capable of binding nucleic acids for treating or
 PT preventing a disease caused by a virus or a disease associated with an
 PT immune response, particularly autoimmune diseases e.g. inflammation or
 PT rheumatoid arthritis -

XX Disclosure; Page 35; 105pp; English.

XX The present invention relates to a new polypeptide capable of binding to
 CC a nucleic acid comprising a receptor nucleotide sequence which can
 CC function as a receptor for virus infection or is involved in an immune
 CC response. The molecules of the invention are useful in the treatment or

CC prevention of a disease caused by a virus or of a disease associated with
 CC an immune response, particularly an autoimmune disease. The zinc finger
 CC polypeptide or a nucleic acid encoding such a polypeptide can be used to
 CC modulate transcription of a receptor nucleotide sequence. The nucleic
 CC acid polypeptide capable of binding to a nucleic acid sequence comprising
 CC a receptor nucleotide sequence, where the receptor is capable of
 CC functioning as a receptor for infection by the virus or is involved in
 CC an immune response, or a nucleic acid encoding such a polypeptide, can
 CC be used in the preparation of a medicament for use in the treatment or
 CC prevention of a disease caused by a virus or of a disease associated
 CC with an immune response. The polypeptides may further be used to treat or
 CC prevent various diseases or syndromes associated with or caused by
 CC malfunction of the receptor as the TNFR1 receptor, such as autoimmune
 CC diseases including inflammation, autoimmune encephalomyelitis, rheumatoid
 CC arthritis or myocarditis. The present amino acid sequence represents a
 CC canonical linker peptide sequence, as described in the invention.

XX
 XX
 SQ Sequence 5 AA;

Query Match 57.1%; Score 4; DB 23; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 GEXP 5
 ||||
 Db 2 GEXP 5

RESULT 45
 AAU76744
 ID AAU76744 standard; Peptide; 5 AA.

XX AAU76744;

DT 21-MAY-2002 (first entry)

DE Nucleic acid binding protein, zinc finger linker peptide #3.

XX Nucleic acid binding protein; zinc finger; enzyme inhibitor;

KW viral replication inhibitor; telomerase inhibitor; gpl20;

KW integrase inhibitor; telomere; G-quadruplex; G-quartet; cytotoxic;

KW human immunodeficiency virus; HIV; acquired immunodeficiency syndrome;

KW AIDS; hyperproliferative disease; cancer; telomerase; apoptosis.

XX Synthetic.

XX WO200204488-A2.

PD 17-JAN-2002.

PF 12-JUL-2001; 2001WO-GB03130.

PR 12-JUL-2000; 2000US-0614679.

XX (SANG-) SANGAMO BIOSCIENCES INC.

PA (UYCA-) UNIV CAMBRIDGE TECH SERVICES LTD.

XX Choo Y, Isalan M, Liu X, Patel S, Balasubramanian S;

DR WPI; 2002-216951/27.

XX Use of a nucleic acid binding polypeptide capable of binding to

PT telomeric, G-quadruplex, or G-quartet nucleic acid as an enzymatic

PT activity inhibitor or cytotoxic agent, for preparing a composition for

PT treating diseases -

XX Disclosure; Page 33; 147pp; English.

XX The invention describes a nucleic acid binding polypeptide (I) capable of
 CC binding to one or more of telomeric, G-quadruplex, or G-quartet nucleic
 CC acid as an inhibitor of enzymatic activity, for the preparation of a
 CC pharmaceutical composition for the treatment of a disease, or as a
 CC cytotoxic agent. (I) is useful for: inhibiting an enzymatic activity;

CC preventing replication of a retrovirus e.g. for treating human
 CC immunodeficiency virus (HIV) infection or acquired immunodeficiency
 CC syndrome (AIDS); treating hyperproliferative disease, such as cancer;
 CC assaying a telomerase activity by providing a nucleic acid substrate for
 CC telomerase; determining the length of a telomere; discriminating between
 CC duplex and quadruplex nucleic acid; detecting telomeric structures in a
 CC system; and identifying a molecule capable of binding to a telomeric,
 CC G-quadruplex, or G-quartet structure in a nucleic acid. (I) is useful for
 CC the preparation of a pharmaceutical composition for the treatment of a
 CC disease, as a cytotoxic agent, and for killing a cell, preferably
 CC by inducing apoptosis. The assay for detecting telomerase activity using
 CC (I) is convenient, rapid, easily automated with liquid handling robotics
 CC and avoids the need to use radioactivity. This sequence represents an
 CC example of a zinc finger linker peptide, used in the engineering of an
 CC nucleic acid binding proteins, described in the method of the invention.

XX
 XX
 SQ Sequence 5 AA;

Query Match 57.1%; Score 4; DB 23; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 GEXP 5
 ||||
 Db 2 GEXP 5

RESULT 46
 ABB05066
 ID ABB05066 standard; Peptide; 5 AA.

XX ABB05066;

DT 27-MAR-2002 (first entry)

DE Linker amino acid sequence #1.

XX Human immunodeficiency virus; HIV; Herpesvirus; HSV; zinc finger;

KW nucleic acid binding protein; viral; promoter; infection; virucide;

KW anti-HIV.

XX Synthetic.

XX WO200185780-A2.

PD 15-NOV-2001.

PF 08-MAY-2001; 2001WO-GB02017.

XX 08-MAY-2000; 2000GB-0011068.

PR 30-MAY-2000; 2000GB-0013106.

PR 02-OCT-2000; 2000WO-GB03765.

PR 19-JAN-2001; 2001GB-0001446.

XX (GEND-) GENDAQ LTD.

PA Choo Y, Demaison C, Moore M, Papworth MA, Reynold L, Ullman CG;

PI Isalan M;

XX WPI; 2002-139420/18.

XX Novel viral nucleic acid binding polypeptide useful for binding viral

PT promoter sequences, and modulating expression of gene linked to viral

PT promoter sequence, and for treating human immunodeficiency virus

PT infection -

XX Disclosure; Page 27; 141pp; English.

XX The present invention describes a polypeptide (I) capable of binding to
 CC a nucleic acid comprising a viral nucleotide sequence. Also described
 CC are: (1) a composition (II) comprising (I) and a excipient, diluent or
 CC carrier; (2) a nucleic acid molecule (III) encoding (I); (3) an
 CC expression vector (IV) comprising (III); (4) a particle (V) harbouring

CC (I), (III) or (IV); and (5) modulating transcription by targeting nucleic acid sequences that overlap with transcription factor binding sites by the use of engineered zinc finger molecules. (I) has virucide and anti-HIV activities. (I) is useful for modulating transcription of a nucleic acid molecule, and for targeting a native viral nucleic acid sequence with a nucleic acid binding polypeptide. (I) is also useful for downregulating a viral function such as viral titre, viral infectivity, viral replication, viral packaging or viral transcription in a cell infected with the virus by contacting the virus and/or the cell with (I). (I) is also useful for modulating a viral function in a system. (I) is highly effective in repressing gene expression from nucleic acid molecules to which they bind. More preferably, they are highly effective in repressing gene expression from the HIV-1 promoter. ABB05064 to ABB05134 and ABA92738 to ABA92779 represent sequences used in the exemplification of the present invention.

XX Sequence 5 AA;
SQ Query Match 57.1%; Score 4; DB 23; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GEXP 5
Db 2 GEXP 5

RESULT 47
ABP57930
ID ABP57930 standard; peptide; 5 AA.

XX AC ABP57930;
XX DT 21-FEB-2003 (first entry)
XX DE Zinc finger protein flexible linker peptide #5.
XX KW Zinc finger; transgenic; transgene; cardiant; cytostatic; vulnery;
XX KW neotropic; neuroprotective; osteopathic; antiarthritic; gene therapy;
XX KW protein therapy; hypertrophic cardiomyopathy; cancer; emphysema; burn;
XX KW Alzheimer's disease; osteoarthritis; linker peptide.
XX OS Synthetic.

XX PN WO200279418-A2.
XX PD 10-OCT-2002.

XX PF 28-MAR-2002; 2002WO-US09703.

XX PR 28-MAR-2001; 2001GB-0007757.

XX PA (SANG-) SANGAMO BIOSCIENCES INC.

XX PI Choo Y, Girdlestone J, Moore M;

XX DR WPI; 2003-103268/09.

XX CC New transgenic animal for studying genetics and diseases in animal models, comprises a nucleic acid binding polypeptide which binds and modulates the expression of a target gene, and a nucleotide sequence encoding the polypeptide -

XX PS Disclosure; Page 39; 90pp; English.

XX CC The invention relates to a novel transgenic non-human animal comprising a heterologous nucleic acid binding polypeptide which binds to a target gene and modulates its expression, where the nucleic acid binding polypeptide is encoded by a transgene, and where the expression of a target gene in at least one cell is modulated compared to a non-transgenic animal. The transgene of the invention has cardiant, cytostatic, vulnery, neotropic, neuroprotective, osteopathic, and antiarthritic activity. The polynucleotide may have a use in gene

CC therapy, and protein therapy. The transgenic animal is useful in studying the relationship between genetics and disease in animal models, and the effect of therapeutic treatments for these diseases. The transgenic animal models are also used for studying developmental processes. The polypeptides and the polynucleotides may be used in treating diseases, such as hypertrophic cardiomyopathy, cancer, emphysema, burns, Alzheimer's disease and osteoarthritis. The present sequence represents a flexible linker peptide, used to join the last residue of the alpha-helix in a zinc finger and the first residue of the beta-sheet in the next zinc finger.

XX SQ Sequence 5 AA;

Query Match 57.1%; Score 4; DB 24; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GEXP 5
Db 2 GEXP 5

RESULT 48
ABP55224
ID ABP55224 standard; Peptide; 5 AA.

XX AC ABP55224;
XX DT 28-JAN-2003 (first entry)
XX DE Linker peptide sequence #6.

XX KW Gene regulation; expression; nucleic acid binding protein; cytostatic; nephrotropic; gene therapy; kidney failure; cancer.

XX OS Synthetic.

XX PN WO200274996-A1.

XX PD 26-SEP-2002.

XX PF 19-MAR-2002; 2002WO-US08554.

XX PR 19-MAR-2001; 2001GB-0006786.

XX PA (SANG-) SANGAMO BIOSCIENCES INC.

XX PI Girdlestone J, England N, Demaison C;

XX DR WPI; 2003-058340/05.

XX CC Regulating expression of a nucleic acid sequence in a primary cell for treating or preventing a disease e.g., cancer, comprises contacting the nucleic acid binding polypeptide with the nucleic acid sequence -

XX PS Disclosure; Page 30; 81pp; English.

XX CC The present invention describes a method (M1) for regulating expression of a nucleic acid sequence in a primary cell comprising providing a nucleic acid binding polypeptide capable of binding to the nucleic acid sequence and contacting the nucleic acid binding polypeptide with the nucleic acid sequence in the primary cell. Also described: (1) a nucleic acid binding polypeptide (I) capable of binding to, and regulating the expression of, a nucleic acid sequence in a primary cell; (2) a primary cell (II) comprising an exogenous nucleic acid binding polypeptide; (3) a pharmaceutical composition (III) comprising the polypeptide or the primary cell and a carrier or diluent; (4) treating (M2) or preventing a disease; or (5) expressing (M3) an exogenous nucleic acid binding polypeptide in a primary cell. (I) has cytostatic and nephrotropic activities, and can be used in gene therapy. The method is useful for treating or preventing a disease e.g., kidney failure or cancer. The present sequence represents a linker peptide which is used in the exemplification of the present invention.

```

XX SQ Sequence 5 AA;
Query Match 57.1%; Score 4; DB 24; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GEKP 5
DB ||||
2 GEKP 5

RESULT 49
ID AAO15222 standard; Peptide; 6 AA.
XX AC AAO15222;
XX DT 05-SEP-2002 (first entry)
XX DE Porphyromonas gingivalis dipeptidylpeptidase-7 (DPP-7) internal region 1.
XX KW Dipeptidylpeptidase-7; DPP-7; enzyme; amidolytic cleavage;
XX KW DPP-7 inhibitor identification; periodontal disease; gingivitis;
XX KW periodontitis.
XX OS Porphyromonas gingivalis.
XX PN WO200238742-A2.
XX PD 16-MAY-2002.
XX PF 08-NOV-2001; 2001WO-US46782.
XX PR 08-NOV-2000; 2000US-246827P.
XX PA (UYGE-) UNIV GEORGIA RES FOUND INC.
XX PI Travis J, Potempa JS, Banbula A, Bugno M;
XX WPI; 2002-490075/52.
XX Novel isolated dipeptidylpeptidase useful for identifying inhibitor of
the dipeptidylpeptidase for protecting an animal from periodontal
disease caused by Porphyromonas gingivalis -
XX Example 6; Page 30; 65pp; English.
XX The invention comprises the amino acid and coding sequence of the
Porphyromonas gingivalis dipeptidylpeptidase-7 (DPP-7) enzyme. The DPP-7
enzyme has amidolytic activity for cleavage of a peptide bond between the
second and third amino acids from the N-terminal end of a target peptide.
The DPP-7 target peptide has an aliphatic or aromatic residue as a
substituent on the alpha-carbon atom of the second amino acid from the
N-terminal end. The DPP-1 protein and DNA sequences of the invention are
useful for identifying a DPP-7 inhibitor. DPP-7 inhibitors are useful for
reducing the growth of a bacterium and protecting an animal from a
periodontal disease caused by Porphyromonas gingivalis (e.g. gingivitis
or periodontitis). The present amino acid sequence represents an internal
region of the Porphyromonas gingivalis dipeptidylpeptidase-7 (DPP-7)
enzyme.
XX SQ Sequence 6 AA;
Query Match 57.1%; Score 4; DB 23; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KPYK 7
DB ||||
3 KPYK 6

```

```

RESULT 50
ID AAR72923 standard; Peptide; 7 AA.
XX AC AAR72923;
XX DT 25-MAR-2003 (updated)
XX DT 29-NOV-1995 (first entry)
XX DE Yeast PPIase cryptic fragment 3.
XX KW Escherichia coli; protein conformation; folding; acceleration;
XX KW PPIase-alpha; peptidyl prolyl cis trans isomerase alpha;
XX KW catalysis; isomerisation; prolyl peptide bond.
XX OS Saccharomyces cerevisiae strain AH22.
XX FH Key Location/Qualifiers
FT Misc-difference 3 /note= "residue not identified in specification"
XX EN EP647714-A1.
XX PD 12-APR-1995.
XX PF 19-JUL-1990; 94EP-0203612.
XX PR 19-JUL-1989; 89JP-0184738.
XX PR 06-OCT-1989; 89JP-0260244.
XX PR 29-DEC-1989; 89JP-0344705.
XX PR 19-JUL-1990; 90EP-0307914.
XX PA (TOFU) TONEN CORP.
XX PI Hayano T, Katou S, Maki N, Suzuki M, Takahashi N;
XX WPI; 1995-140756/19.
XX New E.coli peptidyl prolyl cis trans isomerase beta - used to
accelerate the folding of proteins, partic. for activation of
inactive recombinant proteins
XX Example 2; Page 21; 85pp; English.
XX AAR72921-29 are tryptic fragments of a yeast PPIase (peptidyl prolyl
cis trans isomerase). The yeast PPIase has a single mol. wt. of
about 17 kDa and a single isoelectric point of about 6.2. The enzyme
catalyses the isomerisation of prolyl peptide bonds in proteins and
accelerates the folding of the protein. The inventors are
claiming a PPIase-beta.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX (Updated on 25-MAR-2003 to correct PF field.)
XX (Updated on 25-MAR-2003 to correct PR field.)
XX SQ Sequence 7 AA;
Query Match 57.1%; Score 4; DB 16; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEK 4
DB ||||
4 TGEK 7

RESULT 51
ID AAR72879 standard; Peptide; 7 AA.
XX AC AAR72879;
XX DT 25-MAR-2003 (updated)
XX DT 29-NOV-1995 (first entry)

```

```

XX. Yeast PPIase tryptic fragment 3.
DE
XX
KW Escherichia coli; protein conformation; folding; acceleration;
KW PPIase-alpha; peptidyl prolyl cis trans isomerase alpha;
KW catalysis; isomerisation; prolyl peptide bond.
XX
XX Saccharomyces cerevisiae strain AH22.
OS
XX
XX Key Location/Qualifiers
FH Misc-difference 3
FT /note= "residue not identified in specification"
XX
XX EP647713-A1.
PN
XX
XX 12-APR-1995.
PD
XX
XX 19-JUL-1990; 94EP-0203610.
PF
XX
XX 19-JUL-1989; 89JP-0184738.
PR
XX 06-OCT-1989; 89JP-0260244.
PR
XX 29-DEC-1989; 89JP-0344705.
PR
XX 19-JUL-1990; 90EP-0307914.
PR
XX (TOFU ) TONEN CORP.
PA
XX
XX Hayano T, Katou S, Maki N, Suzuki M, Takahashi N;
PI WPI; 1995-140755/19.
XX
XX New E.coli peptidyl prolyl cis trans isomerase alpha - used to
PT accelerate the folding of proteins, partic. for activation of
PT inactive recombinant proteins
XX
XX Example 2; Page 21; 85pp; English.
PS
XX
XX AAR72877-85 are tryptic fragments of a yeast PPIase (peptidyl prolyl
CC cis trans isomerase). The yeast PPIase has a single mol. wt. of
CC about 17 kDa and a single isoelectric point of about 6.2. The enzyme
CC catalyses the isomerisation of prolyl peptide bonds in proteins and
CC accelerates the folding of the protein. The inventors are
CC claiming the PPIase-alpha.
CC (Updated on 25-MAR-2003 to correct PN field.)
CC (Updated on 25-MAR-2003 to correct PF field.)
CC (Updated on 25-MAR-2003 to correct PR field.)
XX
XX
SQ Sequence 7 AA;
Query Match 57.1%; Score 4; DB 16; Length 7;
Best Local Similarity 100.0%; Pred.No.9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGEK 4
Db ||||
4 TGEK 7
RESULT 52
AAW10599
ID AAW10599 standard; Peptide; 7 AA.
XX
XX AAW10599;
AC
XX
XX 07-MAY-1997 (first entry)
DT
XX
XX Bst DNA polymerase large fragment N-terminal peptide.
DE
XX
XX DNA polymerase; Bst Pol I; thermostable enzyme; exonuclease;
KW DNA sequencing; strand displacement amplification.
XX
XX Bacillus stearothermophilus.
OS
XX
XX EP757100-A1.
PN

```

```

XX 05-FEB-1997.
PD
XX
XX 31-JUL-1996; 96EP-0202169.
PP
XX
XX 02-AUG-1995; 95US-0510215.
PR
XX
XX (NEWE ) NEW ENGLAND BIOLABS INC.
PA
XX
XX Aliotta JM, Kong H, Pelletier JJ;
PI WPI; 1997-111048/11.
XX
XX DNA encoding recombinant thermophilic Bacillus stearothermophilus
PT DNA polymerase I - substantially free of 3'-5', and opt. also 5'-3'
PT exonuclease activity
XX
XX Disclosure; Page 5; 34pp; English.
XX
XX The sequence was detd. of the N-terminal portion (AAW10599) of
CC Bacillus stearothermophilus DNA polymerase (Bst Pol I) large
CC fragment (see also AAW10597), obt'd from native Bst Pol I by
CC subtilisin digestion, and lacking the 3'-5' exonuclease domain of
CC the native enzyme. A PCR primer (AA160826) based on this sequence
CC was used in the PCR amplification of Bst large fragment DNA, using
CC B. stearothermophilus genomic DNA as template. The PCR product was
CC ligated into pMAL-c2 and Bst Pol I large fragment was expressed as
CC a maltose binding protein fusion in E. coli transformants.
XX
XX Sequence 7 AA;
SQ
Query Match 57.1%; Score 4; DB 18; Length 7;
Best Local Similarity 100.0%; Pred.No.9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 GEKP 5
Db ||||
3 GEKP 6
RESULT 53
AAB85411
ID AAB85411 standard; peptide; 7 AA.
XX
XX AAB85411;
AC
XX
XX 17-SEP-2001 (first entry)
DT
XX
XX Peptide linker for generating zinc finger fusion constructs.
DE
XX
XX Nucleic acid binding polypeptide; repressor domain; cardiant; nootropic;
KW circulatory active; anti-inflammatory; dermatological; neuroprotective;
KW cerebroprotective; antibacterial; antifungal; antiviral; antirheumatic;
KW osteopathic; gene therapy; zinc finger; binding site; linker.
XX
XX Synthetic.
OS
XX
XX WO200153480-A1.
PN
XX
XX 26-JUL-2001.
PD
XX
XX 19-JAN-2001; 2001WO-GB00202.
PF
XX
XX 24-JAN-2000; 2000GB-0001582.
PR
XX 30-MAY-2000; 2000GB-0013102.
PR
XX 30-MAY-2000; 2000GB-0013103.
PR
XX 30-MAY-2000; 2000GB-0013104.
PR
XX (GEND-) GENDAQ LTD.
PA
XX
XX Choo Y, Klug A, Moore M;
PI WPI; 2001-451906/48.
XX
XX
DR

```

XX Nucleic acid binding polypeptide, used to identify nucleic acids and
 PT treat inflammatory, neurological, and dermatological disease, comprises
 PT a repressor domain and several nucleic acid binding domains linked by
 PT non-canonical linker(s) -
 XX
 PS Claim 24; Page 105; 142pp; English.
 XX
 CC The invention relates to a nucleic acid (NA) binding polypeptide (I)
 CC comprising a repressor domain and several NA binding domains (BDs) linked
 CC by at least one non-canonical linker. (I) may be used to identify NAs in
 CC a complex mixture, to differentiate single base pair changes in NAs, in
 CC the manufacture of chimeric restriction enzymes, to produce knock out
 CC organisms, and in the treatment of diseases such as: cardiovascular,
 CC inflammatory, metabolic, infectious, neurological, rheumatological,
 CC genetic, dermatological, and musculoskeletal diseases. The invented
 CC methods are used to produce novel NA binding polynucleotides and to
 CC modify existing NA binding polynucleotides comprising several NA BDs.
 CC The novel polypeptide comprises several nucleic acid binding domains
 CC linked by linker sequences. The invented polypeptide is therefore able
 CC to span longer or variable gaps, and a greater number of gaps, between
 CC DNA binding subunits. Sequences AB85403-414 represent peptide linkers
 CC used in the construction of zinc finger fusion constructs.
 XX

SQ Sequence 7 AA;
 Query Match 57.1%; Score 4; DB 22; Length 7;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GEKP 5
 ||||
 Db 4 GEKP 7

RESULT 54

ABG79365

ID ABG79365 standard; Peptide; 7 AA.

XX

AC ABG79365;

XX 15-NOV-2002 (first entry)

XX Canonical linker peptide #7.

XX Receptor; virus infection; immune response; autoimmune disease;

KW TNFRI receptor; inflammation; autoimmune encephalomyelitis;

KW rheumatoid arthritis; myocarditis; canonical linker.

XX Synthetic.

OS

PN WO200257308-A2.

XX

XX 25-JUL-2002.

PD

XX 22-JAN-2002; 2002WO-GB00246.

XX

PF 22-JAN-2001; 2001GB-0001576.

XX

PR 07-FEB-2001; 2001GB-0003032.

XX

XX (SANG-) SANGAMO BIOSCIENCES INC.

PA

XX Moore M, Isalan M, Reynolds L, Ullman C, Girdlestone J;

PI Demaison C, Choo Y;

XX

XX WPI; 2002-590720/63.

DR

XX New polypeptide capable of binding nucleic acids for treating or

PT preventing a disease caused by a virus or a disease associated with an

PT immune response, particularly autoimmune diseases e.g. inflammation or

PT rheumatoid arthritis -

XX Disclosure; Page 35; 105pp; English.

PS

XX The present invention relates to a new polypeptide capable of binding to
 CC a nucleic acid comprising a receptor nucleotide sequence which can
 CC function as a receptor for virus infection or is involved in an immune
 CC response. The molecules of the invention are useful in the treatment or
 CC prevention of a disease caused by a virus or of a disease associated with
 CC an immune response, particularly an autoimmune disease. The zinc finger
 CC polypeptide or a nucleic acid encoding such a polypeptide can be used to
 CC modulate transcription of a receptor nucleotide sequence. The nucleic
 CC acid polypeptide capable of binding to a nucleic acid sequence comprising
 CC a receptor nucleotide sequence, where the receptor is capable of
 CC functioning as a receptor for infection by the virus or is involved in
 CC an immune response, or a nucleic acid encoding such a polypeptide, can
 CC be used in the preparation of a medicament for use in the treatment or
 CC prevention of a disease caused by a virus or of a disease associated
 CC with an immune response. The polypeptides may further be used to treat or
 CC prevent various diseases or syndromes associated with or caused by
 CC malfunction of the receptor as the TNFRI receptor, such as autoimmune
 CC diseases including inflammation, autoimmune encephalomyelitis, rheumatoid
 CC arthritis or myocarditis. The present amino acid sequence represents a
 CC canonical linker peptide sequence, as described in the invention.
 XX

SQ Sequence 7 AA;

Query Match 57.1%; Score 4; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GEKP 5
 ||||
 Db 4 GEKP 7

RESULT 55

AAU76746

ID AAU76746 standard; Peptide; 7 AA.

XX

AC AAU76746;

XX

XX 21-MAY-2002 (first entry)

XX

DE Nucleic acid binding protein, zinc finger linker peptide #5.

XX Nucleic acid binding protein; zinc finger; enzyme inhibitor;

KW viral replication inhibitor; telomerase inhibitor; gp120;

KW integrase inhibitor; telomere; G-quadruplex; G-quartet; cytotoxic;

KW human immunodeficiency virus; HIV; acquired immunodeficiency syndrome;

KW AIDS; hyperproliferative disease; cancer; telomerase; apoptosis.

XX Synthetic.

OS

PN WO200204488-A2.

XX

XX 17-JAN-2002.

PD

XX 12-JUL-2001; 2001WO-GB03130.

XX

XX 12-JUL-2000; 2000US-0614679.

XX

XX (SANG-) SANGAMO BIOSCIENCES INC.

PA (UTCA-) UNIV CAMBRIDGE TECH SERVICES LTD.

XX

XX Choo Y, Isalan M, Liu X, Patel S, Balasubramanian S;

XX WPI; 2002-216951/27.

XX

PT Use of a nucleic acid binding polypeptide capable of binding to

PT telomeric, G-quadruplex, or G-quartet nucleic acid as an enzymatic

PT activity inhibitor or cytotoxic agent, for preparing a composition for

PT treating diseases -

XX Disclosure; Page 33; 147pp; English.

XX

CC The invention describes a nucleic acid binding polypeptide (I) capable of
 CC binding to one or more of telomeric, G-quadruplex, or G-quartet nucleic
 CC acid as an inhibitor of enzymatic activity, for the preparation of a
 CC pharmaceutical composition for the treatment of a disease, or as a
 CC cytotoxic agent. (I) is useful for: inhibiting an enzymatic activity;
 CC preventing replication of a retrovirus e.g. for treating human
 CC immunodeficiency virus (HIV) infection or acquired immunodeficiency
 CC syndrome (AIDS); treating hyperproliferative disease, such as cancer;
 CC assaying a telomerase activity by providing a nucleic acid substrate for
 CC telomerase; determining the length of a telomere; discriminating between
 CC duplex and quadruplex nucleic acid; detecting telomeric structures in a
 CC system; and identifying a molecule capable of binding to a telomeric,
 CC G-quadruplex, or G-quartet structure in a nucleic acid. (I) is useful for
 CC the preparation of a pharmaceutical composition for the treatment of a
 CC disease, as a cytotoxic agent, and for killing a cell, preferably
 CC by inducing apoptosis. The assay for detecting telomerase activity using
 CC (I) is convenient, rapid, easily automated with liquid handling robotics
 CC and avoids the need to use radioactivity. This sequence represents an
 CC example of a zinc finger linker peptide, used in the engineering of
 CC nucleic acid binding proteins, described in the method of the invention.
 XX
 XX
 SQ Sequence 7 AA;

Query Match 57.1%; Score 4; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GEXP 5
 ||||
 DB 4 GEXP 7

RESULT 56
 ABB05068
 ID ABB05068 standard; Peptide; 7 AA.

XX ABB05068;
 DT 27-MAR-2002 (first entry)
 XX
 DE Linker amino acid sequence #3.

XX Human immunodeficiency virus; HIV; Herpesvirus; HSV; zinc finger;
 KW nucleic acid binding protein; viral; promoter; infection; virucide;
 KW anti-HIV.

XX Synthetic.

XX WO200185780-A2.

PN 15-NOV-2001.

PD 08-MAY-2001; 2001WO-GB02017.

XX 08-MAY-2000; 2000GB-0011068.

PR 30-MAY-2000; 2000GB-0013106.

PR 02-OCT-2000; 2000WO-GB03765.

PR 19-JAN-2001; 2001GB-0001446.

XX (GEND-) GENDAQ LTD.

XX Choo Y, Demaison C, Moore M, Papworth MA, Reynold L, Ullman CG;

PI Isalan M;

XX WPI; 2002-139420/18.

DR Novel viral nucleic acid binding polypeptide useful for binding viral

XX promoter sequences, and modulating expression of gene linked to viral

PT promoter sequence, and for treating human immunodeficiency virus

PT infection -

XX Disclosure; Page 27; 141pp; English.

XX

CC The present invention describes a polypeptide (I) capable of binding to
 CC a nucleic acid comprising a viral nucleotide sequence. Also described
 CC are: (1) a composition (II) comprising (I) and a excipient, diluent or
 CC carrier; (2) a nucleic acid molecule (III) encoding (I); (3) an
 CC expression vector (IV) comprising (III); (4) a particle (V) harbouring
 CC (I), (III) or (IV); and (5) modulating transcription by targeting nucleic
 CC acid sequences that overlap with transcription factor binding sites by
 CC the use of engineered zinc finger molecules. (I) has virucide and
 CC anti-HIV activities. (I) is useful for modulating transcription of a
 CC nucleic acid molecule, and for targeting a native viral nucleic acid
 CC sequence with a nucleic acid binding polypeptide. (I) is also useful for
 CC downregulating a viral function such as viral titre, viral infectivity,
 CC viral replication, viral packaging or viral transcription in a cell
 CC infected with the virus by contacting the virus and/or the cell with (I).
 CC (I) is also useful for modulating a viral function in a system. (I) is
 CC highly effective in repressing gene expression from nucleic acid
 CC molecules to which they bind. More preferably, they are highly effective
 CC in repressing gene expression from the HIV-1 promoter. ABB05064 to
 CC ABB05134 and ABA92738 to ABA92779 represent sequences used in the
 CC exemplification of the present invention.
 XX
 XX
 SQ Sequence 7 AA;

Query Match 57.1%; Score 4; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GEXP 5
 ||||
 DB 4 GEXP 7

RESULT 57
 ABP57933
 ID ABP57933 standard; peptide; 7 AA.

XX AC ABP57933;

XX DT 21-FEB-2003 (first entry)

XX Zinc finger protein flexible linker peptide #8.

XX Zinc finger; transgenic; transgene; cardiant; cytostatic; vulnery;
 KW neurotic; neuroprotective; osteopathic; antiarthritic; gene therapy;
 KW protein therapy; hypertrophic cardiomyopathy; cancer; emphysema; burn;
 KW Alzheimer's disease; osteoarthritis; linker peptide.

XX OS Synthetic.

XX WO200279418-A2.

PN 10-OCT-2002.

XX 28-MAR-2002; 2002WO-US09703.

XX 28-MAR-2001; 2001GB-0007757.

XX (SANG-) SANGAMO BIOSCIENCES INC.

XX Choo Y, Girdlestone J, Moore M;

XX WPI; 2003-103268/09.

XX New transgenic animal for studying genetics and diseases in animal
 PT models, comprises a nucleic acid binding polypeptide which binds and
 PT modulates the expression of a target gene, and a nucleotide sequence
 PT encoding the polypeptide -

XX Disclosure; Page 39; 90pp; English.

CC The invention relates to a novel transgenic non-human animal comprising a
 CC heterologous nucleic acid binding polypeptide which binds to a target
 CC gene and modulates its expression, where the nucleic acid binding

CC

CC polypeptide is encoded by a transgene, and where the expression of a
 CC target gene in at least one cell is modulated compared to a
 CC non-transgenic animal. The transgene of the invention has cardiant,
 CC cytostatic, vulnerary, nootropic, neuroprotective, osteopathic, and
 CC antiarthritic activity. The polynucleotide may have a use in gene
 CC therapy, and protein therapy. The transgenic animal is useful in studying
 CC the relationship between genetics and disease in animal models, and the
 CC effect of therapeutic treatments for these diseases. The transgenic
 CC animal models are also used for studying developmental processes. The
 CC polypeptides and the polynucleotides may be used in treating diseases,
 CC such as hypertrophic cardiomyopathy, cancer, emphysema, burns,
 CC Alzheimer's disease and osteoarthritis. The present sequence represents a
 CC flexible linker peptide, used to join the last residue of the alpha-helix
 CC in a zinc finger and the first residue of the beta-sheet in the next zinc
 CC finger.
 XX
 SQ Sequence 7 AA;

Query Match 57.1%; Score 4; DB 24; Length 7;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GEKP 5
 DB 4 GEKP 7

RESULT 58
 ABP55226
 ID ABP55226 standard; Peptide; 7 AA.
 XX
 AC ABP55226;
 XX
 DT 28-JAN-2003 (first entry)
 XX
 DE Linker peptide sequence #8.
 XX
 KW Gene regulation; expression; nucleic acid binding protein; cytostatic;
 KW nephrotropic; gene therapy; kidney failure; cancer.
 XX
 OS Synthetic.
 XX
 XX WO200274996-A1.
 XX
 XX 26-SEP-2002.
 XX
 XX 19-MAR-2002; 2002WO-US09554.
 XX
 XX 19-MAR-2001; 2001GB-0006786.
 XX
 XX (SANG-) SANGAMO BIOSCIENCES INC.
 XX
 XX Girdlestone J, England N, Demaison C;
 XX
 XX WPI; 2003-058340/05.
 XX

PT Regulating expression of a nucleic acid sequence in a primary cell for
 PT treating or preventing a disease e.g., cancer, comprises contacting the
 PT nucleic acid binding polypeptide with the nucleic acid sequence -
 XX
 XX Disclosure; Page 30; 81pp; English.

XX The present invention describes a method (M1) for regulating expression
 CC of a nucleic acid sequence in a primary cell comprising providing a
 CC nucleic acid binding polypeptide capable of binding to the nucleic acid
 CC sequence and contacting the nucleic acid binding polypeptide with the
 CC nucleic acid sequence in the primary cell, also described: (1) a nucleic
 CC acid binding polypeptide (I) capable of binding to, and regulating the
 CC expression of, a nucleic acid sequence in a primary cell; (2) a primary
 CC cell (II) comprising an exogenous nucleic acid binding polypeptide;
 CC (3) a pharmaceutical composition (III) comprising the polypeptide or the
 CC primary cell and a carrier or diluent; (4) treating (M2) or preventing a
 CC disease; or (5) expressing (M3) an exogenous nucleic acid binding

CC polypeptide in a primary cell. (I) has cytostatic and nephrotropic
 CC activities, and can be used in gene therapy. The method is useful for
 CC treating or preventing a disease e.g., kidney failure or cancer. The
 CC present sequence represents a linker peptide which is used in the
 CC exemplification of the present invention.
 XX
 SQ Sequence 7 AA;

Query Match 57.1%; Score 4; DB 24; Length 7;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GEKP 5
 DB 4 GEKP 7

RESULT 59
 AAW74311
 ID AAW74311 standard; peptide; 8 AA.
 XX
 AC AAW74311;
 XX
 DT 05-MAY-1999 (first entry)
 XX
 DE HJ loop peptide K083H102.
 XX

KW HJ loop; serine/threonine kinase; cancer; diabetes; obesity; therapy;
 KW haemorrhagic shock; cardiovascular disease; immunosuppressive disorder;
 KW inflammatory disorder; central nervous system disease; septic shock;
 KW Parkinson's disease; hypertension.
 XX
 CS Synthetic.

XX Key Location/Qualifiers
 PH Modified-site 1 /note= "Myristyl-Gly"
 FT Modified-site 8 /note= "amidated"

XX WO9853051-A1.

XX 26-NOV-1998.

XX 20-MAY-1998; 98WO-US10321.

XX 21-MAY-1997; 97US-0861153.

XX (CHIL-) CHILDRENS MEDICAL CENT.

XX Ben-Sasson SA;

XX WPI; 1999-070143/06.

XX New peptide derivatives for modulating protein tyrosine kinase
 PT activity - comprise a sequence corresponding to the HJ loop of a
 PT protein tyrosine kinase, used for treating cancers or immune
 PT disorders

XX Disclosure; Fig 7; 79pp; English.

XX This sequence represents a peptide of the invention, and is a derivative
 CC of the HJ loop of a serine/threonine kinase (STK). The peptides can be
 CC used for the treatment of disorders caused by overactivity or
 CC underactivity of a STK, e.g. cancer, diabetes, obesity, haemorrhagic
 CC shock, cardiovascular diseases (e.g. thrombosis, atherosclerosis,
 CC arteriosclerosis, cardiac hyper trophy, ischaemia, reperfusion injury and
 CC hypertension), immunosuppressive and inflammatory disorders (e.g. asthma,
 CC psoriasis, systemic lupus erythematosus, diabetes mellitus, suppression
 CC of organ transplant rejection, multiple sclerosis, inflammatory bowel
 CC disease and AIDS), central nervous system diseases (e.g. Alzheimer's
 CC disease, stroke and trauma), septic shock, Parkinson's disease or
 CC hypertension. The peptides can also be used to produce antibodies which

CC can be used to identify cells expressing the STK and to study the
 CC intracellular distribution of the STK. In addition, the peptides can be
 CC used to identify and quantitate ligands which bind the HJ loop of the STK
 CC from which the peptide was derived.

SQ Sequence 8 AA;
 Query Match 57.1%; Score 4; DB 20; Length 8;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KPYK 7
 ||||
 Db 4 KPYK 7

RESULT 60
 AAY98551
 ID AAY98551 standard; Peptide; 9 AA.
 XX AC AAY98551;
 XX DT 31-JUL-2000 (first entry)
 XX DE WT1 derived immunogenic peptide SEQ ID NO:66.
 XX KW WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
 XX KW metastatic disease; mouse; human; Wilm's tumour; immune response;
 XX KW vaccine.
 XX OS Homo sapiens.
 XX PN WO200018795-A2.
 XX PD 06-APR-2000.
 XX PF 30-SEP-1999; 99WO-US22819.
 XX PR 30-SEP-1998; 98US-0164223.
 XX PR 25-MAR-1999; 99US-0276484.
 XX PA (CORI-) CORIXA CORP.
 XX PA (GAIG/) GAIGER A.
 XX PI Gaiger A, Cheever M;
 XX DR WPI; 2000-293107/25.
 XX PT Novel polypeptides comprising an immunogenic portion of a native WT1
 XX PT polypeptide, useful for inhibiting the development of malignant
 XX PT diseases associated with WT1 expression e.g. leukemia or cancer -
 XX PS Claim 4; Page 153; 193pp; English.

XX The present invention describes polypeptides (I) comprising an
 CC immunogenic portion of a native Wilm's Tumour gene product polypeptide,
 CC WT1, (or variants of the immunogenic portion retaining the ability to
 CC react with WT1-specific antisera and/or T-cell lines or clones) and
 CC comprising 16 consecutive amino acids (aa) or less of a native WT1
 CC polypeptide. The polypeptides are useful therapeutically and to
 CC manufacture medicaments for enhancing/inducing an immune response in
 CC patients. The polypeptides, mimetics or polynucleotides can be included
 CC with a carrier/excipient in pharmaceutical compositions or with a
 CC non-specific immune response enhancer (e.g. an adjuvant or enhancer
 CC preferentially enhancing a T cell response) in vaccines. Pharmaceutical
 CC compositions and vaccines can be administered to human patients to
 CC enhance or induce an immune response specific for WT1 or a cell
 CC expressing WT1, useful to inhibit the development of malignant diseases
 CC associated with WT1 expression, e.g. leukemia (especially acute/chronic
 CC myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
 CC breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
 CC AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
 CC AAA13862 represent PCR primers, used in the exemplification of the

CC present invention.
 XX Sequence 9 AA;
 SQ Query Match 57.1%; Score 4; DB 21; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EKPY 6
 ||||
 Db 1 EKPY 4

RESULT 61
 AAY98596
 ID AAY98596 standard; Peptide; 9 AA.
 XX AC AAY98596;
 XX DT 31-JUL-2000 (first entry)
 XX DE WT1 derived immunogenic peptide SEQ ID NO:111.
 XX KW WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
 XX KW metastatic disease; mouse; human; Wilm's tumour; immune response;
 XX KW vaccine.
 XX OS Homo sapiens.
 XX PN WO200018795-A2.
 XX PD 06-APR-2000.
 XX PF 30-SEP-1999; 99WO-US22819.
 XX PR 30-SEP-1998; 98US-0164223.
 XX PR 25-MAR-1999; 99US-0276484.
 XX PA (CORI-) CORIXA CORP.
 XX PA (GAIG/) GAIGER A.
 XX PI Gaiger A, Cheever M;
 XX DR WPI; 2000-293107/25.
 XX PT Novel polypeptides comprising an immunogenic portion of a native WT1
 XX PT polypeptide, useful for inhibiting the development of malignant
 XX PT diseases associated with WT1 expression e.g. leukemia or cancer -
 XX PS Claim 4; Page 160; 193pp; English.
 XX The present invention describes polypeptides (I) comprising an
 CC immunogenic portion of a native Wilm's Tumour gene product polypeptide,
 CC WT1, (or variants of the immunogenic portion retaining the ability to
 CC react with WT1-specific antisera and/or T-cell lines or clones) and
 CC comprising 16 consecutive amino acids (aa) or less of a native WT1
 CC polypeptide. The polypeptides are useful therapeutically and to
 CC manufacture medicaments for enhancing/inducing an immune response in
 CC patients. The polypeptides, mimetics or polynucleotides can be included
 CC with a carrier/excipient in pharmaceutical compositions or with a
 CC non-specific immune response enhancer (e.g. an adjuvant or enhancer
 CC preferentially enhancing a T cell response) in vaccines. Pharmaceutical
 CC compositions and vaccines can be administered to human patients to
 CC enhance or induce an immune response specific for WT1 or a cell
 CC expressing WT1, useful to inhibit the development of malignant diseases
 CC associated with WT1 expression, e.g. leukemia (especially acute/chronic
 CC myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
 CC breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
 CC AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
 CC AAA13862 represent PCR primers, used in the exemplification of the
 XX present invention.
 XX Sequence 9 AA;

```

Query Match          57.1%; Score 4; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEK 4
   ||||
DB 6 TGEK 9

RESULT 62
AAU68650
ID AAU68650 standard; Peptide; 9 AA.
XX
AC AAU68650;
XX
DT 16-JAN-2002 (first entry)
XX
DE Human Wilms' tumour protein, WT1, antigenic peptide #45.
XX
KW Human; Wilms' tumour; WT1; pleural mesothelioma; antigen;
KW leukaemia; acute myeloid leukaemia; AML; chronic myeloid leukaemia; CML;
KW acute lymphocytic leukaemia; ALL; myelodysplastic syndromes;
KW myeloproliferative syndrome; cancer; cytostatic.
XX
OS Homo sapiens.
XX
FN WO200162920-A2.
XX
PD 30-AUG-2001.
XX
PF 22-FEB-2001; 2001WO-US05702.
XX
PR 22-FEB-2000; 2000US-184070P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Cheever MA, Gaiger A;
XX
DR WPI; 2001-648218/74.
XX
PT Composition for the treatment of mesothelioma comprises specific
PT peptides i.e. Wilms' tumour antigen polypeptide derived antigenic
PT fragments -
XX
PS Claim 1; Page 213; 242pp; English.
XX
CC The invention relates to the use of a composition comprising at least a
CC first isolated peptide, of between 9 and 40 amino acids or a first
CC nucleic acid, encoding the peptide, in the manufacture of a medicament
CC for treating or preventing mesothelioma. The peptides are antigenic
CC peptides derived from the Wilms' tumour protein WT1. The composition is
CC useful for the treatment of mesothelioma. Wilms' tumour, preferably
CC pleural mesothelioma and other WT1 associated malignancies e.g.
CC leukaemia (including acute myeloid leukaemia, ALL, and childhood ALL),
CC leukaemia, CML, acute lymphocytic leukaemia, AML, chronic myeloid
CC myelodysplastic syndromes, myeloproliferative syndromes and cancers
CC (e.g. breast, testicular, prostate, lung and ovarian) in mammals,
CC preferably humans. The present sequence is an antigenic peptide of
CC the invention derived from human WT1.
XX
SQ Sequence 9 AA;

Query Match          57.1%; Score 4; DB 22; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EKPX 6
   ||||
DB 1 EKPX 4

RESULT 63

```

```

AAU68695
ID AAU68695 standard; Peptide; 9 AA.
XX
AC AAU68695;
XX
DT 16-JAN-2002 (first entry)
XX
DE Human Wilms' tumour protein, WT1, antigenic peptide #90.
XX
KW Human; Wilms' tumour; WT1; pleural mesothelioma; antigen;
KW leukaemia; acute myeloid leukaemia; AML; chronic myeloid leukaemia; CML;
KW acute lymphocytic leukaemia; ALL; myelodysplastic syndromes;
KW myeloproliferative syndrome; cancer; cytostatic.
XX
OS Homo sapiens.
XX
FN WO200162920-A2.
XX
PD 30-AUG-2001.
XX
PF 22-FEB-2001; 2001WO-US05702.
XX
PR 22-FEB-2000; 2000US-184070P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Cheever MA, Gaiger A;
XX
DR WPI; 2001-648218/74.
XX
PT Composition for the treatment of mesothelioma comprises specific
PT peptides i.e. Wilms' tumour antigen polypeptide derived antigenic
PT fragments -
XX
PS Claim 1; Page 220; 242pp; English.
XX
CC The invention relates to the use of a composition comprising at least a
CC first isolated peptide, of between 9 and 40 amino acids or a first
CC nucleic acid, encoding the peptide, in the manufacture of a medicament
CC for treating or preventing mesothelioma. The peptides are antigenic
CC peptides derived from the Wilms' tumour protein WT1. The composition is
CC useful for the treatment of mesothelioma, Wilms' tumour, preferably
CC pleural mesothelioma and other WT1 associated malignancies e.g.
CC leukaemia (including acute myeloid leukaemia, ALL, and childhood ALL),
CC leukaemia, CML, acute lymphocytic leukaemia, AML, chronic myeloid
CC myelodysplastic syndromes, myeloproliferative syndromes and cancers
CC (e.g. breast, testicular, prostate, lung and ovarian) in mammals,
CC preferably humans. The present sequence is an antigenic peptide of
CC the invention derived from human WT1.
XX
SQ Sequence 9 AA;

Query Match          57.1%; Score 4; DB 22; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEK 4
   ||||
DB 6 TGEK 9

RESULT 64
AAG61883
ID AAG61883 standard; Peptide; 9 AA.
XX
AC AAG61883;
XX
DT 06-JUL-2001 (first entry)
XX
DE Human WT1 immunogenic peptide SEQ ID NO: 66.
XX
KW Human; mouse; immunotherapy; cancer; leukaemia; WT1; Wilms' tumour gene;
KW chromosome 11p13; zinc finger transcription factor.

```

```

XX OS Homo sapiens.
XX PN WO200125273-A2.
XX PD 12-APR-2001.
XX PF 04-OCT-2000; 2000WO-US27465.
XX PR 04-OCT-1999; 99US-0157459.
XX PA (CORI-) CORIXA CORP.
XX PI Skeiky YAW, Xu J, Cheever MA, Reed SG;
XX WI WPI; 2001-328324/34.
XX PT Polypeptide comprising part of the Wilm's Tumour gene product sequence is
XX used in the diagnosis and treatment of malignant diseases e.g. leukemia
XX and cancer associated with WT1 -
XX Claim 4; Page 165; 228pp; English.
XX
XX The present invention describes compositions comprising peptides derived
XX from the Wilm's tumour protein WT1 and methods for their use in treating
XX malignant diseases. Peptides derived from both the murine and human WT1
XX proteins are provided. The human WT1 gene is found on chromosome 11p13,
XX and the protein was shown to be a zinc finger transcription factor. The
XX immunogenic peptides of the invention are particularly useful in the
XX diagnosis and treatment of cancer and leukaemia. The present sequence is
XX a polypeptide described in the exemplification of the invention.
XX
XX Query Match 57.1%; Score 4; DB 22; Length 9;
XX Best Local Similarity 100.0%; Pred. No. 9.3e+05;
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 3 EKPY 6
XX Db 1 EKPY 4
XX
XX RESULT 65
XX AAG61928
XX ID AAG61928 standard; Peptide; 9 AA.
XX AC AAG61928;
XX DT 06-JUL-2001 (first entry)
XX DE Human WT1 immunogenic peptide SEQ ID NO: 111.
XX KW Human; mouse; immunotherapy; cancer; leukaemia; WT1; Wilm's tumour gene;
XX Chromosome 11p13; zinc finger transcription factor.
XX OS Homo sapiens.
XX PN WO200125273-A2.
XX PD 12-APR-2001.
XX PF 04-OCT-2000; 2000WO-US27465.
XX PR 04-OCT-1999; 99US-0157459.
XX PA (CORI-) CORIXA CORP.
XX PI Skeiky YAW, Xu J, Cheever MA, Reed SG;
XX WI WPI; 2001-328324/34.
XX PT Polypeptide comprising part of the Wilm's Tumour gene product sequence is

```

```

PT used in the diagnosis and treatment of malignant diseases e.g. leukemia
PT and cancer associated with WT1 -
XX Claim 4; Page 172; 228pp; English.
XX
XX The present invention describes compositions comprising peptides derived
XX from the Wilm's tumour protein WT1 and methods for their use in treating
XX malignant diseases. Peptides derived from both the murine and human WT1
XX proteins are provided. The human WT1 gene is found on chromosome 11p13,
XX and the protein was shown to be a zinc finger transcription factor. The
XX immunogenic peptides of the invention are particularly useful in the
XX diagnosis and treatment of cancer and leukaemia. The present sequence is
XX a polypeptide described in the exemplification of the invention.
XX
XX Query Match 57.1%; Score 4; DB 22; Length 9;
XX Best Local Similarity 100.0%; Pred. No. 9.3e+05;
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 TGEK 4
XX Db 6 TGEK 9
XX
XX RESULT 66
XX ABJ04613
XX ID ABJ04613 standard; Peptide; 9 AA.
XX AC ABJ04613;
XX DT 24-OCT-2002 (first entry)
XX DE Bone marrow targeting peptide 63.
XX KW BRASIL; targeting peptide; bacterial infection;
XX Biopanning and Rapid Analysis of Selective Interactive Ligands; diabetes;
XX inflammatory arthritis; atherosclerosis; cancer; autoimmune disease;
XX viral infection; cardiovascular disease; degenerative disease.
XX OS Unidentified.
XX PN WO200220822-A2.
XX PD 14-MAR-2002.
XX PF 07-SEP-2001; 2001WO-US28124.
XX PR 08-SEP-2000; 2000US-231266P.
XX PR 17-JAN-2001; 2001US-0765101.
XX PA (TEXA) UNIV TEXAS SYSTEM.
XX PI Arap W, Pasqualini R;
XX WI WPI; 2002-404697/43.
XX
XX Identification of targeting peptides that can be used to treat diseases
XX e.g. cancer and arthritis, by the BRASIL (Biopanning and Rapid Analysis
XX of Selective Ligands) method comprises a single differential
XX centrifugation step -
XX
XX Example 6; Page 80; 167pp; English.
XX
XX The invention comprises a method (BRASIL - Biopanning and Rapid Analysis
XX of Selective Interactive Ligands) to obtain a targeting peptide. The
XX BRASIL method of the invention involves: exposing a target to a phage
XX display library in a first phase; exposing the first phase to a second
XX phase; and separating the phage bound to the target from unbound phage.
XX The BRASIL method of the invention allows cell phages to be separated
XX from the remaining unbound phage in a single differential centrifugation
XX step. When compared to conventional cell panning methods, the BRASIL
XX method shows a significant increase in recovery of specific phage and a

```

CC substantial decrease in background. The BRASIL method is useful for
 CC identifying targeting peptides. The targeting peptides identified by the
 CC method of the invention are useful for treating disease states, such as:
 CC diabetes; inflammatory arthritis; atherosclerosis; cancer; autoimmune
 CC disease; bacterial infection; viral infection; cardiovascular disease and
 CC degenerative disease. The present amino acid sequence represents a
 CC targeting peptide of the invention.

XX Sequence 9 AA;

Query Match 57.1%; Score 4; DB 23; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GEKP 5
 DB 5 GEKP 8

RESULT 67
 ABG33120
 ID ABG33120 standard; Peptide; 9 AA.

XX AC ABG33120;
 XX 15-JUL-2002 (first entry)
 DT DE Human Wt1 immunogenic peptide #43.

XX Human; mouse; cytostatic; immunostimulant; Wt1; cancer;
 KW immune response.
 OS Homo sapiens.

XX WO200228414-A1.

XX 11-APR-2002.
 XX 03-OCT-2001; 2001WO-US31139.

XX 06-OCT-2000; 2000US-0684361.
 XX 09-OCT-2000; 2000US-0685830.
 XX 15-FEB-2001; 2001US-0785019.
 XX 24-AUG-2001; 2001US-0938864.

XX (CORI-) CORIXA CORP.
 PA (GAIG/) GAIGER A.

XX Gaiger A, McNeill PD, Smithgall M, Moulton G, Vedvick TS;
 PI Sleath PR, Mossman S, Evans L, Spies AG, Boydston J;
 XX WPI; 2002-352217/38.

XX Novel isolated Wt1 polynucleotide, and encoded polypeptide, useful for
 PT treating and diagnosing cancer in a patient -
 XX Example 4; Page 175; 260pp; English.

CC The invention relates to an isolated Wt1 polynucleotide (I) and
 CC polypeptide encoded by (I). The Wt1 polynucleotides and polypeptides
 CC are used for treating and detecting cancer in a patient, and for
 CC stimulating an immune response in patient. ABG33070-ABG33405
 CC represent Wt1 amino acid sequences of the invention.

XX Sequence 9 AA;

Query Match 57.1%; Score 4; DB 23; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EKPY 6
 DB 1 EKPY 4

RESULT 68
 ABG33165
 ID ABG33165 standard; Peptide; 9 AA.

XX AC ABG33165;

XX 15-JUL-2002 (first entry)

XX Human Wt1 immunogenic peptide #88.

XX Human; mouse; cytostatic; immunostimulant; Wt1; cancer;
 KW immune response.

XX Homo sapiens.

XX WO200228414-A1.

XX 11-APR-2002.

XX 03-OCT-2001; 2001WO-US31139.

XX 06-OCT-2000; 2000US-0684361.

XX 09-OCT-2000; 2000US-0685830.

XX 15-FEB-2001; 2001US-0785019.

XX 24-AUG-2001; 2001US-0938864.

XX (CORI-) CORIXA CORP.

PA (GAIG/) GAIGER A.

XX Gaiger A, McNeill PD, Smithgall M, Moulton G, Vedvick TS;
 PI Sleath PR, Mossman S, Evans L, Spies AG, Boydston J;

XX WPI; 2002-352217/38.

XX Novel isolated Wt1 polynucleotide, and encoded polypeptide, useful for
 PT treating and diagnosing cancer in a patient -

XX Example 4; Page 182; 260pp; English.

XX The invention relates to an isolated Wt1 polynucleotide (I) and
 CC polypeptide encoded by (I). The Wt1 polynucleotides and polypeptides
 CC are used for treating and detecting cancer in a patient, and for
 CC stimulating an immune response in patient. ABG33070-ABG33405
 CC represent Wt1 amino acid sequences of the invention.

XX Sequence 9 AA;

Query Match 57.1%; Score 4; DB 23; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEK 4
 DB 6 TGEK 9

RESULT 69
 ABRO1871
 ID ABRO1871 standard; Peptide; 9 AA.

XX AC ABRO1871;

XX 19-MAY-2003 (first entry)

XX Human cancer-related protein 74P3B3 HLA peptide #6.

XX Human; cytostatic; vaccine; cancer; immune response; HLA;
 KW human leukocyte antigen.

XX Homo sapiens.

PN WO200283921-A2.
 XX
 PD 24-OCT-2002.
 XX
 XX PF 10-APR-2002; 2002WO-US11654.
 XX
 PR 10-APR-2001; 2001US-282739P.
 PR 10-APR-2001; 2001US-283112P.
 PR 25-APR-2001; 2001US-286630P.
 XX
 XX PA (AGEN-) AGENSYS INC.

XX Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
 PI Morrison K, Morrison RK, Raitano AB;
 PI WPI; 2003-075555/07.
 DR

XX New composition comprising a substance that modulates the structure of
 PT proteins and polynucleotides, useful for therapeutic, prognostic and
 PT diagnostic reagents for eliciting cellular or humoral immune response
 PT in cancer patients -
 XX

PS Claim 13; Page 131; 102lpp; English.
 XX

XX The present invention relates to novel human cancer-related genes and
 CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
 CC proteins are useful for eliciting a humoral or cellular immune response.
 CC The genes are useful as probes and primers for the amplification and/or
 CC detection of genes, mRNAs or their fragments, as reagents for the
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of
 CC directing the expression of the protein, as tools for modulating or
 CC inhibiting the expression of genes and/or translation of transcripts, and
 CC as therapeutic agents. The proteins and peptides are useful as
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
 CC from the invention.
 XX

SQ Sequence 9 AA;

Query Match 57.1%; Score 4; DB 24; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EKPY 6
 ||||
 Db 3 EKPY 6

RESULT 70
 ABR01897
 ID ABR01897 standard; Peptide; 9 AA.

XX ABR01897;
 XX
 XX 19-MAY-2003 (first entry)
 DT
 XX Human cancer-related protein 74P3B3 HLA peptide #32.
 DE
 XX Human; cytostatic; vaccine; cancer; immune response; HLA;
 KW human leukocyte antigen.
 KW
 XX Homo sapiens.
 OS
 XX WO200283921-A2.
 FN
 XX 24-OCT-2002.
 PD
 XX 10-APR-2002; 2002WO-US11654.

PF 10-APR-2001; 2001US-282739P.
 PR 10-APR-2001; 2001US-283112P.
 PR 25-APR-2001; 2001US-286630P.
 XX

PA (AGEN-) AGENSYS INC.
 XX
 PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
 PI Morrison K, Morrison RK, Raitano AB;
 XX
 DR WPI; 2003-075555/07.

XX New composition comprising a substance that modulates the structure of
 PT proteins and polynucleotides, useful for therapeutic, prognostic and
 PT diagnostic reagents for eliciting cellular or humoral immune response
 PT in cancer patients -
 XX

PS Claim 13; Page 131; 102lpp; English.
 XX

XX The present invention relates to novel human cancer-related genes and
 CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
 CC proteins are useful for eliciting a humoral or cellular immune response.
 CC The genes are useful as probes and primers for the amplification and/or
 CC detection of genes, mRNAs or their fragments, as reagents for the
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of
 CC directing the expression of the protein, as tools for modulating or
 CC inhibiting the expression of genes and/or translation of transcripts, and
 CC as therapeutic agents. The proteins and peptides are useful as
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
 CC from the invention.
 XX

SQ Sequence 9 AA;

Query Match 57.1%; Score 4; DB 24; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EKPY 6
 ||||
 Db 6 EKPY 9

RESULT 71
 ABR02120

ID ABR02120 standard; Peptide; 9 AA.

XX ABR02120;
 AC
 XX 19-MAY-2003 (first entry)
 DT
 XX Human cancer-related protein 74P3B3 HLA peptide #255.
 DE
 XX Human; cytostatic; vaccine; cancer; immune response; HLA;
 KW human leukocyte antigen.
 KW
 XX Homo sapiens.
 OS
 XX WO200283921-A2.
 FN
 XX 24-OCT-2002.
 PD
 XX 10-APR-2002; 2002WO-US11654.

PF 10-APR-2001; 2001US-282739P.
 PR 10-APR-2001; 2001US-283112P.
 PR 25-APR-2001; 2001US-286630P.
 XX
 XX PA (AGEN-) AGENSYS INC.

XX Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
 PI Morrison K, Morrison RK, Raitano AB;
 PI WPI; 2003-075555/07.
 DR

XX New composition comprising a substance that modulates the structure of
 PT proteins and polynucleotides, useful for therapeutic, prognostic and
 PT diagnostic reagents for eliciting cellular or humoral immune response
 PT in cancer patients -

PT in cancer patients -

XX Claim 13; Page 133; 1021pp; English.

XX The present invention relates to novel human cancer-related genes and

CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and

CC proteins are useful for eliciting a humoral or cellular immune response.

CC The genes are useful as probes and primers for the amplification and/or

CC detection of genes, mRNAs or their fragments, as reagents for the

CC diagnosis and/or prognosis of cancer, as coding sequences capable of

CC directing the expression of the protein, as tools for modulating or

CC inhibiting the expression of genes and/or translation of transcripts, and

CC as therapeutic agents. The proteins and peptides are useful as

CC therapeutic, prognostic and diagnostic reagents for cancer. The present

CC sequence is a human leukocyte antigen (HLA) peptide, used in an example

XX from the invention.

SQ Sequence 9 AA;

Query Match 57.1%; Score 4; DB 24; Length 9;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 EKPX 6

Db |||||

2 EKPX 5

RESULT 72

ABR02125

ID ABR02125 standard; Peptide; 9 AA.

AC ABR02125;

XX 19-MAY-2003 (first entry)

XX Human cancer-related protein 74P3B3 HLA peptide #260.

DE Human; cytostatic; vaccine; cancer; immune response; HLA;

KW human leukocyte antigen.

KW Homo sapiens.

OS WO200283921-A2.

PN 24-OCT-2002.

PD 10-APR-2002; 2002WO-US11654.

XX 10-APR-2001; 2001US-282739P.

PR 10-APR-2001; 2001US-283112P.

PR 25-APR-2001; 2001US-286630P.

XX (AGEN-) AGENSYS INC.

PA Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;

PI Morrison K, Morrison RK, Raitano AB;

XX WPI; 2003-075555/07.

XX New composition comprising a substance that modulates the structure of

PT proteins and polynucleotides, useful for therapeutic, prognostic and

PT diagnostic reagents for eliciting cellular or humoral immune response

PT in cancer patients -

XX Claim 13; Page 133; 1021pp; English.

XX The present invention relates to novel human cancer-related genes and

CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and

CC proteins are useful for eliciting a humoral or cellular immune response.

CC The genes are useful as probes and primers for the amplification and/or

CC detection of genes, mRNAs or their fragments, as reagents for the

CC diagnosis and/or prognosis of cancer, as coding sequences capable of

CC directing the expression of the protein, as tools for modulating or

CC inhibiting the expression of genes and/or translation of transcripts, and

CC as therapeutic agents. The proteins and peptides are useful as

CC therapeutic, prognostic and diagnostic reagents for cancer. The present

CC sequence is a human leukocyte antigen (HLA) peptide, used in an example

XX from the invention.

CC directing the expression of the protein, as tools for modulating or

CC inhibiting the expression of genes and/or translation of transcripts, and

CC as therapeutic agents. The proteins and peptides are useful as

CC therapeutic, prognostic and diagnostic reagents for cancer. The present

CC sequence is a human leukocyte antigen (HLA) peptide, used in an example

XX from the invention.

XX Sequence 9 AA;

Query Match 57.1%; Score 4; DB 24; Length 9;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 EKPX 6

Db |||||

5 EKPX 8

RESULT 73

ABR02147

ID ABR02147 standard; Peptide; 9 AA.

XX ABR02147;

XX 19-MAY-2003 (first entry)

XX Human cancer-related protein 74P3B3 HLA peptide #282.

DE Human; cytostatic; vaccine; cancer; immune response; HLA;

KW human leukocyte antigen.

KW Homo sapiens.

OS WO200283921-A2.

PN 24-OCT-2002.

PD 10-APR-2002; 2002WO-US11654.

XX 10-APR-2001; 2001US-282739P.

PR 10-APR-2001; 2001US-283112P.

PR 25-APR-2001; 2001US-286630P.

XX (AGEN-) AGENSYS INC.

PA Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;

PI Morrison K, Morrison RK, Raitano AB;

XX WPI; 2003-075555/07.

XX New composition comprising a substance that modulates the structure of

PT proteins and polynucleotides, useful for therapeutic, prognostic and

PT diagnostic reagents for eliciting cellular or humoral immune response

PT in cancer patients -

XX Claim 13; Page 133; 1021pp; English.

XX The present invention relates to novel human cancer-related genes and

CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and

CC proteins are useful for eliciting a humoral or cellular immune response.

CC The genes are useful as probes and primers for the amplification and/or

CC detection of genes, mRNAs or their fragments, as reagents for the

CC diagnosis and/or prognosis of cancer, as coding sequences capable of

CC directing the expression of the protein, as tools for modulating or

CC inhibiting the expression of genes and/or translation of transcripts, and

CC as therapeutic agents. The proteins and peptides are useful as

CC therapeutic, prognostic and diagnostic reagents for cancer. The present

CC sequence is a human leukocyte antigen (HLA) peptide, used in an example

XX from the invention.

XX Sequence 9 AA;

Query Match 57.1%; Score 4; DB 24; Length 9;

Best Local Similarity 100.0%; Pred. No. 9.3e+05; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EKP 6
Db 6 EKP 9

RESULT 74

ABR02153
ID ABR02153 standard; Peptide; 9 AA.

XX AC ABR02153;

XX DT 19-MAY-2003 (first entry)

XX DE Human cancer-related protein 74P3B3 HLA peptide #288.

XX KW Human; cytostatic; vaccine; cancer; immune response; HLA;

XX KW human leukocyte antigen.

XX OS Homo sapiens.

XX PN WO200283921-A2.

XX PD 24-OCT-2002.

XX PF 10-APR-2002; 2002WO-US11654.

XX PR 10-APR-2001; 2001US-282739P.

XX PR 10-APR-2001; 2001US-283112P.

XX PR 25-APR-2001; 2001US-286630P.

XX PA (AGEN-) AGENSYS INC.

XX PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;

XX PI Morrison K, Morrison RK, Raitano AB;

XX DR WPI; 2003-075555/07.

XX PT New composition comprising a substance that modulates the structure of proteins and polynucleotides, useful for therapeutic, prognostic and diagnostic reagents for eliciting cellular or humoral immune response in cancer patients

XX PS Claim 13; Page 133; 1021pp; English.

XX CC The present invention relates to novel human cancer-related genes and proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and proteins are useful for eliciting a humoral or cellular immune response. The genes are useful as probes and primers for the amplification and/or detection of genes, mRNAs or their fragments, as reagents for the diagnosis and/or prognosis of cancer, as coding sequences capable of directing the expression of the protein, as tools for modulating or inhibiting the expression of genes and/or translation of transcripts, and as therapeutic agents. The proteins and peptides are useful as therapeutic, prognostic and diagnostic reagents for cancer. The present sequence is a human leukocyte antigen (HLA) peptide, used in an example from the invention.

XX SQ Sequence 9 AA;

Query Match 57.1%; Score 4; DB 24; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EKP 6
Db 3 EKP 6

RESULT 75

ABR02158

ABR02158 standard; Peptide; 9 AA.

XX AC ABR02158;

XX DT 19-MAY-2003 (first entry)

XX DE Human cancer-related protein 74P3B3 HLA peptide #293.

XX KW Human; cytostatic; vaccine; cancer; immune response; HLA;

XX KW human leukocyte antigen.

XX OS Homo sapiens.

XX PN WO200283921-A2.

XX PD 24-OCT-2002.

XX PF 10-APR-2002; 2002WO-US11654.

XX PR 10-APR-2001; 2001US-282739P.

XX PR 10-APR-2001; 2001US-283112P.

XX PR 25-APR-2001; 2001US-286630P.

XX PA (AGEN-) AGENSYS INC.

XX PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;

XX PI Morrison K, Morrison RK, Raitano AB;

XX DR WPI; 2003-075555/07.

XX PT New composition comprising a substance that modulates the structure of proteins and polynucleotides, useful for therapeutic, prognostic and diagnostic reagents for eliciting cellular or humoral immune response in cancer patients

XX PS Claim 13; Page 133; 1021pp; English.

XX CC The present invention relates to novel human cancer-related genes and proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and proteins are useful for eliciting a humoral or cellular immune response. The genes are useful as probes and primers for the amplification and/or detection of genes, mRNAs or their fragments, as reagents for the diagnosis and/or prognosis of cancer, as coding sequences capable of directing the expression of the protein, as tools for modulating or inhibiting the expression of genes and/or translation of transcripts, and as therapeutic agents. The proteins and peptides are useful as therapeutic, prognostic and diagnostic reagents for cancer. The present sequence is a human leukocyte antigen (HLA) peptide, used in an example from the invention.

XX SQ Sequence 9 AA;

Query Match 57.1%; Score 4; DB 24; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EKP 6
Db 1 EKP 4

RESULT 76

ABR02271

ID ABR02271 standard; Peptide; 9 AA.

XX AC ABR02271;

XX DT 19-MAY-2003 (first entry)

XX DE Human cancer-related protein 74P3B3 HLA peptide #406.

XX KW Human; cytostatic; vaccine; cancer; immune response; HLA;

XX KW human leukocyte antigen.

```

XX OS Homo sapiens.
XX PN WO200283921-A2.
XX PD 24-OCT-2002.
XX PF 10-APR-2002; 2002WO-US11654.
XX PR 10-APR-2001; 2001US-282739P.
XX PR 10-APR-2001; 2001US-283112P.
XX PR 25-APR-2001; 2001US-286630P.
XX PA (AGEN-) AGENSYS INC.
XX PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
XX PI Morrison K, Morrison RK, Raitano AB;
XX DR WPI; 2003-075555/07.
XX PT New composition comprising a substance that modulates the structure of
XX PT proteins and polynucleotides, useful for therapeutic, prognostic and
XX PT diagnostic reagents for eliciting cellular or humoral immune response
XX PT in cancer patients.
XX PS Claim 13; Page 135; 1021pp; English.
XX CC The present invention relates to novel human cancer-related genes and
XX CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
XX CC proteins are useful for eliciting a humoral or cellular immune response.
XX CC The genes are useful as probes and primers for the amplification and/or
XX CC detection of genes, mRNAs or their fragments, as reagents for the
XX CC diagnosis and/or prognosis of cancer, as coding sequences capable of
XX CC directing the expression of the protein, as tools for modulating or
XX CC inhibiting the expression of genes and/or translation of transcripts, and
XX CC as therapeutic agents. The proteins and peptides are useful as
XX CC therapeutic, prognostic and diagnostic reagents for cancer. The present
XX CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
XX CC from the invention.
XX SQ Sequence 9 AA;
XX Query Match 57.1%; Score 4; DB 24; Length 9;
XX Best Local Similarity 100.0%; Pred. No. 9.3e+05;
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 EKPX 6
DB 3 EKPX 6
RESULT 77
ABR02301
ID ABR02301 standard; Peptide; 9 AA.
XX AC ABR02301;
XX DT 19-MAY-2003 (first entry)
XX DE Human cancer-related protein 74P3B3 HLA peptide #436.
XX KW Human; cytostatic; vaccine; cancer; immune response; HLA;
XX KW human leukocyte antigen.
XX OS Homo sapiens.
XX PN WO200283921-A2.
XX PD 24-OCT-2002.
XX PR 10-APR-2002; 2002WO-US11654.
XX PR 10-APR-2001; 2001US-282739P.
XX PR 25-APR-2001; 2001US-286630P.
XX PA (AGEN-) AGENSYS INC.
XX PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
XX PI Morrison K, Morrison RK, Raitano AB;
XX DR WPI; 2003-075555/07.
XX PR 10-APR-2001; 2001US-282739P.

```

```

PR 10-APR-2001; 2001US-283112P.
PR 25-APR-2001; 2001US-286630P.
XX PA (AGEN-) AGENSYS INC.
XX PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
XX PI Morrison K, Morrison RK, Raitano AB;
XX DR WPI; 2003-075555/07.
XX PT New composition comprising a substance that modulates the structure of
XX PT proteins and polynucleotides, useful for therapeutic, prognostic and
XX PT diagnostic reagents for eliciting cellular or humoral immune response
XX PT in cancer patients.
XX PS Claim 13; Page 135; 1021pp; English.
XX CC The present invention relates to novel human cancer-related genes and
XX CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
XX CC proteins are useful for eliciting a humoral or cellular immune response.
XX CC The genes are useful as probes and primers for the amplification and/or
XX CC detection of genes, mRNAs or their fragments, as reagents for the
XX CC diagnosis and/or prognosis of cancer, as coding sequences capable of
XX CC directing the expression of the protein, as tools for modulating or
XX CC inhibiting the expression of genes and/or translation of transcripts, and
XX CC as therapeutic agents. The proteins and peptides are useful as
XX CC therapeutic, prognostic and diagnostic reagents for cancer. The present
XX CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
XX CC from the invention.
XX SQ Sequence 9 AA;
XX Query Match 57.1%; Score 4; DB 24; Length 9;
XX Best Local Similarity 100.0%; Pred. No. 9.3e+05;
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 EKPX 6
DB 6 EKPX 9
RESULT 78
ABR02331
ID ABR02331 standard; Peptide; 9 AA.
XX AC ABR02331;
XX DT 19-MAY-2003 (first entry)
XX DE Human cancer-related protein 74P3B3 HLA peptide #466.
XX KW Human; cytostatic; vaccine; cancer; immune response; HLA;
XX KW human leukocyte antigen.
XX OS Homo sapiens.
XX PN WO200283921-A2.
XX PD 24-OCT-2002.
XX PR 10-APR-2002; 2002WO-US11654.
XX PR 10-APR-2001; 2001US-282739P.
XX PR 10-APR-2001; 2001US-283112P.
XX PR 25-APR-2001; 2001US-286630P.
XX PA (AGEN-) AGENSYS INC.
XX PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
XX PI Morrison K, Morrison RK, Raitano AB;
XX DR WPI; 2003-075555/07.
XX PR 10-APR-2001; 2001US-282739P.

```

PT New composition comprising a substance that modulates the structure of
 PT proteins and polynucleotides, useful for therapeutic, prognostic and
 PT diagnostic reagents for eliciting cellular or humoral immune response
 PT in cancer patients -
 XX
 PS Claim 13; Page 135; 102lpp; English.
 XX
 CC The present invention relates to novel human cancer-related genes and
 CC proteins (AB278120-AB278168 and ABR01789-ABR01861). The genes and
 CC proteins are useful for eliciting a humoral or cellular immune response.
 CC The genes are useful as probes and primers for the amplification and/or
 CC detection of genes, mRNAs or their fragments, as reagents for the
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of
 CC directing the expression of the protein, as tools for modulating or
 CC inhibiting the expression of the protein, as tools for modulating or
 CC as therapeutic agents. The proteins and peptides are useful as
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
 CC from the invention.
 XX
 SQ Sequence 9 AA;

Query Match 57.1%; Score 4; DB 24; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 EXPY 6
 Db 2 EXPY 5

RESULT 79
 ABR02346
 ID ABR02346 standard; Peptide; 9 AA.

AC ABR02346;

DT 19-MAY-2003 (first entry)

DE Human cancer-related protein 74P3B3 HLA peptide #481.

KW Human; cytostatic; vaccine; cancer; immune response; HLA;
 KW human leukocyte antigen.

OS Homo sapiens.

PN WO200283921-A2.

PD 24-OCT-2002.

PF 10-APR-2002; 2002WO-US11654.

PR 10-APR-2001; 2001US-282739P.

PR 10-APR-2001; 2001US-283112P.

PR 25-APR-2001; 2001US-286630P.

XX (AGEN-) AGENSYS INC.

XX Jakobovits A, Challita-Eid PM, Paris M, Ge W, Hubert RS;

XX Morrison K, Morrison RK, Raitano AB;

XX WPI; 2003-075555/07.

XX New composition comprising a substance that modulates the structure of
 XX proteins and polynucleotides, useful for therapeutic, prognostic and
 XX diagnostic reagents for eliciting cellular or humoral immune response
 XX in cancer patients -

XX Claim 13; Page 135; 102lpp; English.

XX The present invention relates to novel human cancer-related genes and
 CC proteins (AB278120-AB278168 and ABR01789-ABR01861). The genes and
 CC proteins are useful for eliciting a humoral or cellular immune response.

CC The genes are useful as probes and primers for the amplification and/or
 CC detection of genes, mRNAs or their fragments, as reagents for the
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of
 CC directing the expression of the protein, as tools for modulating or
 CC inhibiting the expression of genes and/or translation of transcripts, and
 CC as therapeutic agents. The proteins and peptides are useful as
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
 CC from the invention.

SQ Sequence 9 AA;

Query Match 57.1%; Score 4; DB 24; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 EXPY 6

Db 5 EXPY 8

RESULT 80

ABR02470

ID ABR02470 standard; Peptide; 9 AA.

AC ABR02470;

DT 19-MAY-2003 (first entry)

DE Human cancer-related protein 74P3B3 HLA peptide #605.

KW Human; cytostatic; vaccine; cancer; immune response; HLA;
 KW human leukocyte antigen.

OS Homo sapiens.

PN WO200283921-A2.

PD 24-OCT-2002.

PF 10-APR-2002; 2002WO-US11654.

PR 10-APR-2001; 2001US-282739P.

PR 10-APR-2001; 2001US-283112P.

PR 25-APR-2001; 2001US-286630P.

XX (AGEN-) AGENSYS INC.

XX Jakobovits A, Challita-Eid PM, Paris M, Ge W, Hubert RS;

XX Morrison K, Morrison RK, Raitano AB;

XX WPI; 2003-075555/07.

XX New composition comprising a substance that modulates the structure of
 XX proteins and polynucleotides, useful for therapeutic, prognostic and
 XX diagnostic reagents for eliciting cellular or humoral immune response
 XX in cancer patients -

XX Claim 13; Page 137; 102lpp; English.

XX The present invention relates to novel human cancer-related genes and
 CC proteins (AB278120-AB278168 and ABR01789-ABR01861). The genes and
 CC proteins are useful for eliciting a humoral or cellular immune response.
 CC The genes are useful as probes and primers for the amplification and/or
 CC detection of genes, mRNAs or their fragments, as reagents for the
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of
 CC directing the expression of the protein, as tools for modulating or
 CC inhibiting the expression of genes and/or translation of transcripts, and
 CC as therapeutic agents. The proteins and peptides are useful as
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
 CC from the invention.

SQ Sequence 9 AA;

Query Match 57.1%; Score 4; DB 24; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EKP 6
 ||||
 Db 3 EKP 6

RESULT 81

ABR02514
 ID ABR02514 standard; Peptide; 9 AA.

XX ABR02514;
 AC
 XX
 DT 19-MAY-2003 (first entry)
 XX
 DE Human cancer-related protein 74P3B3 HLA peptide #649.
 XX
 KW Human; cytostatic; vaccine; cancer; immune response; HLA;
 KW human leukocyte antigen.
 XX
 OS Homo sapiens.
 XX
 FN WO200283921-A2.
 XX
 PD 24-OCT-2002.
 XX
 PF 10-APR-2002; 2002WO-US11654.
 XX
 PR 10-APR-2001; 2001US-282739P.
 XX
 PR 10-APR-2001; 2001US-283112P.
 XX
 PR 25-APR-2001; 2001US-286630P.
 XX
 PA (AGEN-) AGENSYS INC.
 XX
 PI Jakobovits A, Challita-Bid PM, Paris M, Ge W, Hubert RS;
 PI Morrison K, Morrison RK, Raitano AB;
 XX
 DR WPI; 2003-075555/07.
 XX
 PT New composition comprising a substance that modulates the structure of
 PT proteins and polynucleotides, useful for therapeutic, prognostic and
 PT diagnostic reagents for eliciting cellular or humoral immune response
 PT in cancer patients -
 XX
 PS Claim 13; Page 137; 1021pp; English.
 XX
 CC The present invention relates to novel human cancer-related genes and
 CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
 CC proteins are useful for eliciting a humoral or cellular immune response.
 CC The genes are useful as probes and primers for the amplification and/or
 CC detection of genes, mRNAs or their fragments, as reagents for the
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of
 CC directing the expression of the protein, as tools for modulating or
 CC inhibiting the expression of genes and/or translation of transcripts, and
 CC as therapeutic agents. The proteins and peptides are useful as
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
 CC from the invention.

SQ Sequence 9 AA;

Query Match 57.1%; Score 4; DB 24; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EKP 6
 ||||
 Db 5 EKP 8

RESULT 83

ABR02698
 ID ABR02698 standard; Peptide; 9 AA.

XX ABR02698;
 AC
 XX
 DT 19-MAY-2003 (first entry)
 XX
 DE Human cancer-related protein 74P3B3 HLA peptide #833.

RESULT 82

ABR02529
 ID ABR02529 standard; Peptide; 9 AA.

XX ABR02529;
 AC
 XX
 DT 19-MAY-2003 (first entry)
 XX
 DE Human cancer-related protein 74P3B3 HLA peptide #664.
 XX
 KW Human; cytostatic; vaccine; cancer; immune response; HLA;
 KW human leukocyte antigen.
 XX
 OS Homo sapiens.
 XX
 FN WO200283921-A2.
 XX
 PD 24-OCT-2002.
 XX
 PF 10-APR-2002; 2002WO-US11654.
 XX
 PR 10-APR-2001; 2001US-282739P.
 XX
 PR 10-APR-2001; 2001US-283112P.
 XX
 PR 25-APR-2001; 2001US-286630P.
 XX
 PA (AGEN-) AGENSYS INC.
 XX
 PI Jakobovits A, Challita-Bid PM, Paris M, Ge W, Hubert RS;
 PI Morrison K, Morrison RK, Raitano AB;
 XX
 DR WPI; 2003-075555/07.
 XX
 PT New composition comprising a substance that modulates the structure of
 PT proteins and polynucleotides, useful for therapeutic, prognostic and
 PT diagnostic reagents for eliciting cellular or humoral immune response
 PT in cancer patients -
 XX
 PS Claim 13; Page 137; 1021pp; English.
 XX
 CC The present invention relates to novel human cancer-related genes and
 CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
 CC proteins are useful for eliciting a humoral or cellular immune response.
 CC The genes are useful as probes and primers for the amplification and/or
 CC detection of genes, mRNAs or their fragments, as reagents for the
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of
 CC directing the expression of the protein, as tools for modulating or
 CC inhibiting the expression of genes and/or translation of transcripts, and
 CC as therapeutic agents. The proteins and peptides are useful as
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
 CC from the invention.

SQ Sequence 9 AA;

Query Match 57.1%; Score 4; DB 24; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EKP 6
 ||||
 Db 6 EKP 9

XX Human; cytostatic; vaccine; cancer; immune response; HLA;
 KW human leukocyte antigen.
 XX Homo sapiens.
 OS WO200283921-A2.
 PN 24-OCT-2002.
 XX 10-APR-2002; 2002WO-US11654.
 XX 10-APR-2001; 2001US-282739P.
 PR 10-APR-2001; 2001US-283112P.
 PR 25-APR-2001; 2001US-286630P.
 XX (AGEN-) AGENSYS INC.
 PA Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
 PI Morrison K, Morrison RK, Raitano AB;
 XX WPI; 2003-075555/07.
 XX New composition comprising a substance that modulates the structure of
 PT proteins and polynucleotides, useful for therapeutic, prognostic and
 PT diagnostic reagents for eliciting cellular or humoral immune response
 PT in cancer patients -
 XX Claim 13; Page 139; 1021pp; English.
 XX The present invention relates to novel human cancer-related genes and
 CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
 CC proteins are useful for eliciting a humoral or cellular immune response.
 CC The genes are useful as probes and primers for the amplification and/or
 CC detection of genes, mRNAs or their fragments, as reagents for the
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of
 CC directing the expression of the protein, as tools for modulating or
 CC inhibiting the expression of genes and/or translation of transcripts, and
 CC as therapeutic agents. The proteins and peptides are useful as
 CC therapeutic agents. The proteins and peptides are useful as
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
 CC from the invention.
 XX Sequence 9 AA;
 SQ Query Match 57.1%; Score 4; DB 24; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 EKPY 6
 DB 1 EKPY 4
 RESULT 84
 ABR02699
 ID ABR02699 standard; Peptide; 9 AA.
 AC ABR02699;
 XX 19-MAY-2003 (first entry)
 XX Human cancer-related protein 74P3B3 HLA peptide #834.
 DE Human; cytostatic; vaccine; cancer; immune response; HLA;
 KW human leukocyte antigen.
 XX Homo sapiens.
 OS WO200283921-A2.
 PN 24-OCT-2002.
 XX 10-APR-2002; 2002WO-US11654.
 XX 10-APR-2001; 2001US-282739P.
 PR 10-APR-2001; 2001US-283112P.
 PR 25-APR-2001; 2001US-286630P.
 XX (AGEN-) AGENSYS INC.
 PA Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
 PI Morrison K, Morrison RK, Raitano AB;
 XX WPI; 2003-075555/07.
 XX New composition comprising a substance that modulates the structure of
 PT proteins and polynucleotides, useful for therapeutic, prognostic and
 PT diagnostic reagents for eliciting cellular or humoral immune response
 PT in cancer patients -
 XX Claim 13; Page 139; 1021pp; English.
 XX The present invention relates to novel human cancer-related genes and
 CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
 CC proteins are useful for eliciting a humoral or cellular immune response.
 CC The genes are useful as probes and primers for the amplification and/or
 CC detection of genes, mRNAs or their fragments, as reagents for the
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of
 CC directing the expression of the protein, as tools for modulating or
 CC inhibiting the expression of genes and/or translation of transcripts, and
 CC as therapeutic agents. The proteins and peptides are useful as
 CC therapeutic agents. The proteins and peptides are useful as
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
 CC from the invention.
 XX Sequence 9 AA;
 SQ Query Match 57.1%; Score 4; DB 24; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 EKPY 6
 DB 1 EKPY 4

PF 10-APR-2002; 2002WO-US11654.
 XX 10-APR-2001; 2001US-282739P.
 PR 10-APR-2001; 2001US-283112P.
 PR 25-APR-2001; 2001US-286630P.
 XX (AGEN-) AGENSYS INC.
 PA Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
 PI Morrison K, Morrison RK, Raitano AB;
 XX WPI; 2003-075555/07.
 XX New composition comprising a substance that modulates the structure of
 PT proteins and polynucleotides, useful for therapeutic, prognostic and
 PT diagnostic reagents for eliciting cellular or humoral immune response
 PT in cancer patients -
 XX Claim 13; Page 139; 1021pp; English.
 XX The present invention relates to novel human cancer-related genes and
 CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
 CC proteins are useful for eliciting a humoral or cellular immune response.
 CC The genes are useful as probes and primers for the amplification and/or
 CC detection of genes, mRNAs or their fragments, as reagents for the
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of
 CC directing the expression of the protein, as tools for modulating or
 CC inhibiting the expression of genes and/or translation of transcripts, and
 CC as therapeutic agents. The proteins and peptides are useful as
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
 CC from the invention.
 XX Sequence 9 AA;
 SQ Query Match 57.1%; Score 4; DB 24; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 EKPY 6
 DB 5 EKPY 8
 RESULT 85
 ABR02705
 ID ABR02705 standard; Peptide; 9 AA.
 AC ABR02705;
 XX 19-MAY-2003 (first entry)
 XX Human cancer-related protein 74P3B3 HLA peptide #840.
 DE Human; cytostatic; vaccine; cancer; immune response; HLA;
 KW human leukocyte antigen.
 XX Homo sapiens.
 OS WO200283921-A2.
 PN 24-OCT-2002.
 XX 10-APR-2002; 2002WO-US11654.
 XX 10-APR-2001; 2001US-282739P.
 PR 10-APR-2001; 2001US-283112P.
 PR 25-APR-2001; 2001US-286630P.
 XX (AGEN-) AGENSYS INC.
 PA Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
 PI Morrison K, Morrison RK, Raitano AB;
 XX WPI; 2003-075555/07.

XX WPI; 2003-075555/07.
 XX
 PT New composition comprising a substance that modulates the structure of
 PT proteins and polynucleotides, useful for therapeutic, prognostic and
 PT diagnostic reagents for eliciting cellular or humoral immune response
 PT in cancer patients -
 XX
 XX Claim 13; Page 139; 1021pp; English.
 PS
 CC The present invention relates to novel human cancer-related genes and
 CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
 CC proteins are useful for eliciting a humoral or cellular immune response.
 CC The genes are useful as probes and primers for the amplification and/or
 CC detection of genes, mRNAs or their fragments, as reagents for the
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of
 CC directing the expression of the protein, as tools for modulating or
 CC inhibiting the expression of genes and/or translation of transcripts, and
 CC as therapeutic agents. The proteins and peptides are useful as
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
 CC from the invention.
 XX
 SQ Sequence 9 AA;
 Query Match 57.1%; Score 4; DB 24; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 EKP 6
 DB |||||
 6 EKP 9
 RESULT 86
 ABR02746
 ID ABR02746 standard; Peptide; 9 AA.
 AC ABR02746;
 XX
 DT 19-MAY-2003 (first entry)
 XX
 DE Human cancer-related protein 74P3B3 HLA peptide #881.
 XX
 KW Human; cytostatic; vaccine; cancer; immune response; HLA;
 KW human leukocyte antigen.
 XX
 OS Homo sapiens.
 XX
 PN WO200283921-A2.
 XX
 PD 24-OCT-2002.
 XX
 PF 10-APR-2002; 2002WO-US11654.
 XX
 PR 10-APR-2001; 2001US-282739P.
 PR 10-APR-2001; 2001US-283112P.
 PR 25-APR-2001; 2001US-286630P.
 XX
 PA (AGEN-) AGENSYS INC.
 XX
 PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
 PI Morrison K, Morrison RK, Raitano AB;
 XX
 DR WPI; 2003-075555/07.
 XX
 PT New composition comprising a substance that modulates the structure of
 PT proteins and polynucleotides, useful for therapeutic, prognostic and
 PT diagnostic reagents for eliciting cellular or humoral immune response
 PT in cancer patients -
 XX
 XX Claim 13; Page 139; 1021pp; English.
 PS
 CC The present invention relates to novel human cancer-related genes and
 CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
 CC proteins are useful for eliciting a humoral or cellular immune response.
 CC The genes are useful as probes and primers for the amplification and/or
 CC detection of genes, mRNAs or their fragments, as reagents for the
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of
 CC directing the expression of the protein, as tools for modulating or
 CC inhibiting the expression of genes and/or translation of transcripts, and
 CC as therapeutic agents. The proteins and peptides are useful as
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
 CC from the invention.
 XX
 SQ Sequence 9 AA;
 Query Match 57.1%; Score 4; DB 24; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 EKP 6
 DB |||||
 6 EKP 9

CC The present invention relates to novel human cancer-related genes and
 CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
 CC proteins are useful for eliciting a humoral or cellular immune response.
 CC The genes are useful as probes and primers for the amplification and/or
 CC detection of genes, mRNAs or their fragments, as reagents for the
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of
 CC directing the expression of the protein, as tools for modulating or
 CC inhibiting the expression of genes and/or translation of transcripts, and
 CC as therapeutic agents. The proteins and peptides are useful as
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
 CC from the invention.
 XX
 SQ Sequence 9 AA;
 Query Match 57.1%; Score 4; DB 24; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 EKP 6
 DB |||||
 3 EKP 6
 RESULT 87
 ABR02866
 ID ABR02866 standard; Peptide; 9 AA.
 XX
 AC ABR02866;
 XX
 DT 19-MAY-2003 (first entry)
 XX
 DE Human cancer-related protein 74P3B3 HLA peptide #1001.
 XX
 KW Human; cytostatic; vaccine; cancer; immune response; HLA;
 KW human leukocyte antigen.
 XX
 OS Homo sapiens.
 XX
 PN WO200283921-A2.
 XX
 PD 24-OCT-2002.
 XX
 PF 10-APR-2002; 2002WO-US11654.
 XX
 PR 10-APR-2001; 2001US-282739P.
 PR 10-APR-2001; 2001US-283112P.
 PR 25-APR-2001; 2001US-286630P.
 XX
 PA (AGEN-) AGENSYS INC.
 XX
 PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
 PI Morrison K, Morrison RK, Raitano AB;
 XX
 DR WPI; 2003-075555/07.
 XX
 PT New composition comprising a substance that modulates the structure of
 PT proteins and polynucleotides, useful for therapeutic, prognostic and
 PT diagnostic reagents for eliciting cellular or humoral immune response
 PT in cancer patients -
 XX
 XX Claim 13; Page 141; 1021pp; English.
 PS
 CC The present invention relates to novel human cancer-related genes and
 CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
 CC proteins are useful for eliciting a humoral or cellular immune response.
 CC The genes are useful as probes and primers for the amplification and/or
 CC detection of genes, mRNAs or their fragments, as reagents for the
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of
 CC directing the expression of the protein, as tools for modulating or
 CC inhibiting the expression of genes and/or translation of transcripts, and
 CC as therapeutic agents. The proteins and peptides are useful as
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present

CC* sequence is a human leukocyte antigen (HLA) peptide, used in an example
 CC from the invention.

XX Sequence 9 AA;

Query Match 57.1%; Score 4; DB 24; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EXPY 6
 ||||
 Db 5 EXPY 8

RESULT 88

ABR02889
 ID ABR02889 standard; Peptide; 9 AA.

XX ABR02889;

DT 19-MAY-2003 (first entry)

DE Human cancer-related protein 74P3B3 HLA peptide #1024.

XX Human; cytostatic; vaccine; cancer; immune response; HLA;
 KW human leukocyte antigen.

XX Homo sapiens.

XX WO200283921-A2.

XX 24-OCT-2002.

XX 10-APR-2002; 2002WO-US11654.

PR 10-APR-2001; 2001US-282739P.

PR 10-APR-2001; 2001US-283112P.

PR 25-APR-2001; 2001US-286630P.

XX (AGEN-) AGENSYS INC.

PI Jakobovits A, Challita-Eid PM, Paris M, Ge W, Hubert RS;

PI Morrison K, Morrison RK, Raitano AB;

XX WPI; 2003-075555/07.

XX New composition comprising a substance that modulates the structure of
 PT proteins and polynucleotides, useful for therapeutic, prognostic and
 PT diagnostic reagents for eliciting cellular or humoral immune response
 PT in cancer patients -

XX Claim 13; Page 141; 1021pp; English.

XX The present invention relates to novel human cancer-related genes and
 CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
 CC proteins are useful for eliciting a humoral or cellular immune response.
 CC The genes are useful as probes and primers for the amplification and/or
 CC detection of genes, mRNAs or their fragments, as reagents for the
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of
 CC directing the expression of the protein, as tools for modulating or
 CC inhibiting the expression of genes and/or translation of transcripts, and
 CC as therapeutic agents. The proteins and peptides are useful as
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
 CC from the invention.

XX Sequence 9 AA;

Query Match 57.1%; Score 4; DB 24; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EXPY 6

Db ||||
 1 EXPY 4

RESULT 89

ABR02902
 ID ABR02902 standard; Peptide; 9 AA.

XX ABR02902;

DT 19-MAY-2003 (first entry)

DE Human cancer-related protein 74P3B3 HLA peptide #1037.

XX Human; cytostatic; vaccine; cancer; immune response; HLA;
 KW human leukocyte antigen.

XX Homo sapiens.

XX WO200283921-A2.

XX 24-OCT-2002.

XX 10-APR-2002; 2002WO-US11654.

PR 10-APR-2001; 2001US-282739P.

PR 10-APR-2001; 2001US-283112P.

PR 25-APR-2001; 2001US-286630P.

XX (AGEN-) AGENSYS INC.

PI Jakobovits A, Challita-Eid PM, Paris M, Ge W, Hubert RS;

PI Morrison K, Morrison RK, Raitano AB;

XX WPI; 2003-075555/07.

XX New composition comprising a substance that modulates the structure of
 PT proteins and polynucleotides, useful for therapeutic, prognostic and
 PT diagnostic reagents for eliciting cellular or humoral immune response
 PT in cancer patients -

XX Claim 13; Page 141; 1021pp; English.

XX The present invention relates to novel human cancer-related genes and
 CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
 CC proteins are useful for eliciting a humoral or cellular immune response.
 CC The genes are useful as probes and primers for the amplification and/or
 CC detection of genes, mRNAs or their fragments, as reagents for the
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of
 CC directing the expression of the protein, as tools for modulating or
 CC inhibiting the expression of genes and/or translation of transcripts, and
 CC as therapeutic agents. The proteins and peptides are useful as
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
 CC from the invention.

XX Sequence 9 AA;

Query Match 57.1%; Score 4; DB 24; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EXPY 6

Db ||||
 6 EXPY 9

RESULT 90

ABR03066
 ID ABR03066 standard; Peptide; 9 AA.

XX ABR03066;

XX

XX
PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
PI Morrison K, Morrison RK, Raitano AB;
XX WPI; 2003-075555/07.
XX
XX New composition comprising a substance that modulates the structure of
PT proteins and polynucleotides, useful for therapeutic, prognostic and
PT diagnostic reagents for eliciting cellular or humoral immune response
PT in cancer patients -
XX
XX Claim 13; Page 143; 1021pp; English.
XX
XX The present invention relates to novel human cancer-related genes and
CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
CC proteins are useful for eliciting a humoral or cellular immune response.
CC The genes are useful as probes and primers for the amplification and/or
CC detection of genes, mRNAs or their fragments, as reagents for the
CC diagnosis and/or prognosis of cancer, as coding sequences capable of
CC directing the expression of the protein, as tools for modulating or
CC inhibiting the expression of genes and/or translation of transcripts, and
CC as therapeutic agents. The proteins and peptides are useful as
CC therapeutic, prognostic and diagnostic reagents for cancer. The present
CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
CC from the invention.
XX
SQ Sequence 9 AA;

Query Match 57.1%; Score 4; DB 24; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 EXPY 6
Db 1 EXPY 4

RESULT 93
ABR03153
ID ABR03153 standard; Peptide; 9 AA.

AC ABR03153;

DT 19-MAY-2003 (first entry)

DE Human cancer-related protein 74P3B3 HLA peptide #1288.

KW Human; cytostatic; vaccine; cancer; immune response; HLA;
KW human leukocyte antigen.

OS Homo sapiens.

FN WO200283921-A2.

PD 24-OCT-2002.

PF 10-APR-2002; 2002WO-US11654.

PR 10-APR-2001; 2001US-282739P.

PR 10-APR-2001; 2001US-283112P.

PR 25-APR-2001; 2001US-286630P.

PA (AGEN-) AGENSYS INC.

PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;

PI Morrison K, Morrison RK, Raitano AB;

DR WPI; 2003-075555/07.

XX New composition comprising a substance that modulates the structure of
PT proteins and polynucleotides, useful for therapeutic, prognostic and
PT diagnostic reagents for eliciting cellular or humoral immune response
PT in cancer patients -

XX
PS Claim 13; Page 143; 1021pp; English.
XX
XX The present invention relates to novel human cancer-related genes and
CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
CC proteins are useful for eliciting a humoral or cellular immune response.
CC The genes are useful as probes and primers for the amplification and/or
CC detection of genes, mRNAs or their fragments, as reagents for the
CC diagnosis and/or prognosis of cancer, as coding sequences capable of
CC directing the expression of the protein, as tools for modulating or
CC inhibiting the expression of genes and/or translation of transcripts, and
CC as therapeutic agents. The proteins and peptides are useful as
CC therapeutic, prognostic and diagnostic reagents for cancer. The present
CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
CC from the invention.
XX
SQ Sequence 9 AA;

Query Match 57.1%; Score 4; DB 24; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 EXPY 6
Db 2 EXPY 5

RESULT 94

ABR05079

ID ABR05079 standard; Peptide; 9 AA.

AC ABR05079;

DT 19-MAY-2003 (first entry)

DE Human cancer-related protein 83P4B8 HLA peptide #414.

KW Human; cytostatic; vaccine; cancer; immune response; HLA;
KW human leukocyte antigen.

OS Homo sapiens.

FN WO200283921-A2.

PD 24-OCT-2002.

PF 10-APR-2002; 2002WO-US11654.

PR 10-APR-2001; 2001US-282739P.

PR 10-APR-2001; 2001US-283112P.

PR 25-APR-2001; 2001US-286630P.

PA (AGEN-) AGENSYS INC.

PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;

PI Morrison K, Morrison RK, Raitano AB;

DR WPI; 2003-075555/07.

XX New composition comprising a substance that modulates the structure of
PT proteins and polynucleotides, useful for therapeutic, prognostic and
PT diagnostic reagents for eliciting cellular or humoral immune response
PT in cancer patients -

PS Claim 13; Page 163; 1021pp; English.

XX The present invention relates to novel human cancer-related genes and
CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
CC proteins are useful for eliciting a humoral or cellular immune response.
CC The genes are useful as probes and primers for the amplification and/or
CC detection of genes, mRNAs or their fragments, as reagents for the
CC diagnosis and/or prognosis of cancer, as coding sequences capable of
CC directing the expression of the protein, as tools for modulating or

CC inhibiting the expression of genes and/or translation of transcripts, and
 CC as therapeutic agents. The proteins and peptides are useful as
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
 CC from the invention.

XX Sequence 9 AA;

Query Match 57.1%; Score 4; DB 24; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEK 4
 ||||
 Db 5 TGEK 8

RESULT 95
 ABR05312
 ID ABR05312 standard; Peptide; 9 AA.

XX AC ABR05312;

XX DT 19-MAY-2003 (first entry)

XX DE Human cancer-related protein 83P4B8 HLA peptide #647.

XX KW Human; cytostatic; vaccine; cancer; immune response; HLA;
 human leukocyte antigen.

XX OS Homo sapiens.

XX PN WO200283921-A2.

XX PD 24-OCT-2002.

XX PF 10-APR-2002; 2002WO-US11654.

XX PR 10-APR-2001; 2001US-282739P.

XX PR 10-APR-2001; 2001US-283112P.

XX PR 25-APR-2001; 2001US-286630P.

XX PA (AGEN-) AGENSYS INC.

XX PI Jakobovits A, Challita-Eid PM, Paris M, Ge W, Hubert RS;

XX PI Morrison K, Morrison RK, Raitano AB;

XX PF WPI; 2003-075555/07.

XX PT New composition comprising a substance that modulates the structure of
 proteins and polynucleotides, useful for therapeutic, prognostic and
 PT diagnostic reagents for eliciting cellular or humoral immune response
 PT in cancer patients -

XX PS Claim 13; Page 165; 1021pp; English.

XX The present invention relates to novel human cancer-related genes and
 CC proteins (AB278120-AB278168 and ABR01789-ABR01861). The genes and
 CC proteins are useful for eliciting a humoral or cellular immune response.
 CC The genes are useful as probes and primers for the amplification and/or
 CC detection of genes, mRNAs or their fragments, as reagents for the
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of
 CC directing the expression of the protein, as tools for modulating or
 CC inhibiting the expression of genes and/or translation of transcripts, and
 CC as therapeutic agents. The proteins and peptides are useful as
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
 CC from the invention.

XX Sequence 9 AA;

Query Match 57.1%; Score 4; DB 24; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TGEK 4
 ||||
 Db 5 TGEK 8

RESULT 96

ABR05327
 ID ABR05327 standard; Peptide; 9 AA.

XX AC ABR05327;

XX DT 19-MAY-2003 (first entry)

XX DE Human cancer-related protein 83P4B8 HLA peptide #662.

XX KW Human; cytostatic; vaccine; cancer; immune response; HLA;
 human leukocyte antigen.

XX OS Homo sapiens.

XX PN WO200283921-A2.

XX PD 24-OCT-2002.

XX PF 10-APR-2002; 2002WO-US11654.

XX PR 10-APR-2001; 2001US-282739P.

XX PR 10-APR-2001; 2001US-283112P.

XX PR 25-APR-2001; 2001US-286630P.

XX PA (AGEN-) AGENSYS INC.

XX PI Jakobovits A, Challita-Eid PM, Paris M, Ge W, Hubert RS;

XX PI Morrison K, Morrison RK, Raitano AB;

XX PF WPI; 2003-075555/07.

XX PT New composition comprising a substance that modulates the structure of
 proteins and polynucleotides, useful for therapeutic, prognostic and
 PT diagnostic reagents for eliciting cellular or humoral immune response
 PT in cancer patients -

XX PS Claim 13; Page 165; 1021pp; English.

XX The present invention relates to novel human cancer-related genes and
 CC proteins (AB278120-AB278168 and ABR01789-ABR01861). The genes and
 CC proteins are useful for eliciting a humoral or cellular immune response.
 CC The genes are useful as probes and primers for the amplification and/or
 CC detection of genes, mRNAs or their fragments, as reagents for the
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of
 CC directing the expression of the protein, as tools for modulating or
 CC inhibiting the expression of genes and/or translation of transcripts, and
 CC as therapeutic agents. The proteins and peptides are useful as
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
 CC from the invention.

XX Sequence 9 AA;

Query Match 57.1%; Score 4; DB 24; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGEK 4
 ||||
 Db 3 TGEK 6

RESULT 97

ABR05353

ID ABR05353 standard; Peptide; 9 AA.

KW ABR05353;
 AC
 XX
 DT
 XX
 DE 19-MAY-2003 (first entry)
 XX
 DE Human cancer-related protein 83P4B8 HLA peptide #688.
 XX
 KW Human; cytostatic; vaccine; cancer; immune response; HLA;
 KW human leukocyte antigen.
 XX
 OS Homo sapiens.
 XX
 PN WO200283921-A2.
 XX
 XX
 PD 24-OCT-2002.
 XX
 PF 10-APR-2002; 2002WO-US11654.
 XX
 XX
 PR 10-APR-2001; 2001US-282739P.
 PR 10-APR-2001; 2001US-283112P.
 PR 25-APR-2001; 2001US-286630P.
 XX
 PA (AGEN-) AGENSYS INC.
 XX
 XX Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
 PI Morrison K, Morrison RK, Raitano AB;
 XX
 DR WPI; 2003-075555/07.
 XX
 XX New composition comprising a substance that modulates the structure of
 PT proteins and polynucleotides, useful for therapeutic, prognostic and
 PT diagnostic reagents for eliciting cellular or humoral immune response
 PT in cancer patients -
 XX
 PS Claim 13; Page 165; 1021pp; English.
 XX
 CC The present invention relates to novel human cancer-related genes and
 CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
 CC proteins are useful for eliciting a humoral or cellular immune response.
 CC The genes are useful as probes and primers for the amplification and/or
 CC detection of genes, mRNAs or their fragments, as reagents for the
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of
 CC directing the expression of the protein, as tools for modulating or
 CC inhibiting the expression of genes and/or translation of transcripts, and
 CC as therapeutic agents. The proteins and peptides are useful as
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
 CC from the invention.
 XX
 SQ Sequence 9 AA;
 Query Match 57.1%; Score 4; DB 24; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 TGEK 4
 Db 6 TGEK 9
 RESULT 98
 AAU05603
 ID AAU05603 standard; Peptide; 10 AA.
 XX
 AC AAU05603;
 XX
 DT 24-OCT-2001 (first entry)
 XX
 DE N-terminal peptide addition 27 from library B.
 XX
 KW Glucocerebrosidase; GCB; lysosomal storage disease;
 KW Gaucher's disease; Fabry's disease; Farber's disease;
 KW G_m_1 gangliosidosis; Tay-Sach's disease; Niemann-Pick disease;

KW Shindler disease; Hunter syndrome; Sly syndrome; Huler syndrome;
 KW Scheie syndrome; N-terminal peptide addition; library B.
 XX
 OS Synthetic.
 XX
 PN WO200149830-A2.
 XX
 PD 12-JUL-2001.
 XX
 PF 29-DEC-2000; 2000WO-DK00743.
 XX
 XX 30-DEC-1999; 99DK-0001891.
 PR 02-JUN-2000; 2000DK-0000865.
 PR 02-JUN-2000; 2000DK-0000866.
 PR 30-JUN-2000; 2000DK-0001027.
 XX
 PA (MAXY-) MAXYGEN APS.
 XX
 XX Okkels JS, Jensen AD, Halkier T, Jensen RB, Schambye HT;
 PI WPI; 2001-465259/50.
 DR
 XX Improved lysosomal enzymes and lysosomal enzyme activators useful for
 PT treating Gaucher's disease -
 XX
 PS Example 3; Page 62; 97pp; English.
 XX
 CC The sequence represents an N-terminal peptide addition (from library B)
 CC engineered onto the human lysosomal enzyme glucocerebrosidase,
 CC GCB. GCB is the enzyme involved in Gaucher's disease, a lysosomal
 CC storage disease. The invention relates to introducing new glycosylation
 CC sites into lysosomal enzymes/activators like GCB to improve their
 CC bioactivity. The novel polypeptides are used for the prevention and
 CC treatment of Gaucher's disease, Fabry's disease, Farber's disease,
 CC G_m_1 gangliosidosis, Tay-Sach's disease, Niemann-Pick disease, Shindler
 CC disease, Hunter syndrome, Sly syndrome, Huler and Huler/Scheie
 CC syndromes.
 XX
 SQ Sequence 10 AA;
 Query Match 57.1%; Score 4; DB 22; Length 10;
 Best Local Similarity 100.0%; Pred. No. 1.17e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 4 KPYK 7
 Db 2 KPYK 5
 RESULT 99
 AAB85413
 ID AAB85413 standard; peptide; 10 AA.
 XX
 AC AAB85413;
 XX
 DT 17-SEP-2001 (first entry)
 XX
 DE Peptide linker for generating zinc finger fusion constructs.
 XX
 KW Nucleic acid binding polypeptide; repressor domain; cardiant; nootropic;
 KW circulatory active; anti-inflammatory; dermatological; neuroprotective;
 KW cerebroprotective; antibacterial; antifungal; antiviral; antirheumatic;
 KW osteopathic; gene therapy; zinc finger; binding site; linker.
 XX
 OS Synthetic.
 XX
 PN WO200153480-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 19-JAN-2001; 2001WO-GB00202.
 XX
 XX 24-JAN-2000; 2000GB-0001582.

PR 30-MAY-2000; 2000GB-0013102.
PR 30-MAY-2000; 2000GB-0013103.
PR 30-MAY-2000; 2000GB-0013104.

XX (GENB-) GENDAQ LTD.

XX Choo Y, Klug A, Moore M;

XX WPI; 2001-451906/48.

XX Nucleic acid binding polypeptide, used to identify nucleic acids and
PT treat inflammatory, neurological, and dermatological disease, comprises
PT a repressor domain and several nucleic acid binding domains linked by
PT non-canonical linker(s) -

XX Claim 24; Page 105; 142pp; English.

XX The invention relates to a nucleic acid (NA) binding polypeptide (I)
CC comprising a repressor domain and several NA binding domains (BDs) linked
CC by at least one non-canonical linker. (I) may be used to identify NAs in
CC a complex mixture, to differentiate single base pair changes in NAs, in
CC the manufacture of chimeric restriction enzymes, to produce knock out
CC organisms, and in the treatment of diseases such as: cardiovascular,
CC inflammatory, metabolic, infectious, neurological, rheumatological,
CC genetic, dermatological, and musculoskeletal diseases. The invented
CC methods are used to produce novel NA binding polynucleotides and to
CC modify existing NA binding polynucleotides comprising several NA BDs.
CC The novel polypeptide comprises several nucleic acid binding domains
CC linked by linker sequences. The invented polypeptide is therefore able
CC to span longer or variable gaps, and a greater number of gaps, between
CC DNA binding subunits. Sequences AAB85405-414 represent peptide linkers
CC used in the construction of zinc finger fusion constructs.

XX Sequence 10 AA;

Query Match 57.1%; Score 4; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GEKP 5
Db ||||
7 GEKP 10

RESULT 100

ABG79367

ID ABG79367 standard; Peptide; 10 AA.

XX AC

XX ABG79367;

XX 15-NOV-2002 (first entry)

XX Canonical linker peptide #9.

XX Receptor; virus infection; immune response; autoimmune disease;
KW TNFRI receptor; inflammation; autoimmune encephalomyelitis;
KW rheumatoid arthritis; myocarditis; canonical linker.

XX Synthetic.

XX WO200257308-A2.

XX 25-JUL-2002.

XX 22-JAN-2002; 2002WO-GB00246.

XX 22-JAN-2001; 2001GB-0001576.

PR 07-FEB-2001; 2001GB-0003032.

XX (SANG-) SANGAMO BIOSCIENCES INC.

XX Moore M, Isalan M, Reynolds L, Ullman C, Girdlestone J;
PI Demaison C, Choo Y;

XX WPI; 2002-590720/63.

XX New polypeptide capable of binding nucleic acids for treating or
PT preventing a disease caused by a virus or a disease associated with an
PT immune response, particularly autoimmune diseases e.g. inflammation or
PT rheumatoid arthritis -

XX Disclosure; Page 35; 105pp; English.

XX The present invention relates to a new polypeptide capable of binding to
CC a nucleic acid comprising a receptor nucleotide sequence which can
CC function as a receptor for virus infection or is involved in an immune
CC response. The molecules of the invention are useful in the treatment or
CC prevention of a disease caused by a virus or of a disease associated with
CC an immune response, particularly an autoimmune disease. The zinc finger
CC polypeptide or a nucleic acid encoding such a polypeptide can be used to
CC modulate transcription of a receptor nucleotide sequence. The nucleic
CC acid polypeptide capable of binding to a nucleic acid sequence comprising
CC a receptor nucleotide sequence, where the receptor is capable of
CC functioning as a receptor for infection by the virus or is involved in
CC an immune response, or a nucleic acid encoding such a polypeptide, can
CC be used in the preparation of a medicament for use in the treatment or
CC prevention of a disease caused by a virus or of a disease associated
CC with an immune response. The polypeptides may further be used to treat or
CC prevent various diseases or syndromes associated with or caused by
CC malfunction of the receptor as the TNFRI receptor, such as autoimmune
CC diseases including inflammation, autoimmune encephalomyelitis, rheumatoid
CC arthritis or myocarditis. The present amino acid sequence represents a
CC canonical linker peptide sequence, as described in the invention.

XX Sequence 10 AA;

Query Match 57.1%; Score 4; DB 23; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GEKP 5
Db ||||
7 GEKP 10

Search completed: October 21, 2003, 18:23:48
Job time : 63.8889 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2003 CompuGen Ltd.
 OM protein - protein search, using sw model
 Run on: October 21, 2003, 18:19:26 ; Search time 55.6111 Seconds
 (without alignments)
 32.482 Million cell updates/sec

Title: US-10-057-890A-8
 Perfect score: 7
 Sequence: 1 TGEKPYK 7

Scoring table: OLIGO
 Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1349

Minimum DB seq length: 0
 Maximum DB seq length: 10

Post-processing: Listing first 150 summaries

Database : SPTREMBL_23:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phase:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rvivirus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3	42.9	8	11 P70243	P70243 mus musculus
2	3	42.9	8	11 Q62528	Q62528 mus spretus
3	3	42.9	9	2 Q9E214	Q9E214 sodalis glo
4	3	42.9	9	4 Q9R735	Q9R735 streptomyce
5	3	42.9	9	4 Q16605	Q16605 homo sapien
6	3	42.9	9	11 Q9QZ48	Q9QZ48 mus musculus
7	2	28.6	7	2 Q47505	Q47505 escherichia
8	2	28.6	7	12 Q66205	Q66205 transmissib
9	2	28.6	8	2 Q9R772	Q9R772 escherichia
10	2	28.6	8	2 Q9AGP4	Q9AGP4 arthrobacte
11	2	28.6	8	2 Q4453	Q4453 rhizobiales
12	2	28.6	8	4 Q16428	Q16428 homo sapien
13	2	28.6	8	4 Q15888	Q15888 homo sapien
14	2	28.6	8	4 Q9UJ50	Q9UJ50 homo sapien
15	2	28.6	8	5 Q9TWH6	Q9TWH6 perinereis
16	2	28.6	8	5 Q94695	Q94695 physarum po

17	2	28.6	8	6	Q9TRX8	Q9TRX8 bos taurus
18	2	28.6	8	6	Q9XSY1	Q9XSY1 canis famil
19	2	28.6	8	6	Q9MYL5	Q9MYL5 pongo pygma
20	2	28.6	8	6	Q95M23	Q95M23 sus scrofa
21	2	28.6	8	7	Q95213	Q95213 oryctolagus
22	2	28.6	8	8	Q94PX5	Q94PX5 felis silve
23	2	28.6	8	8	Q94EX7	Q94EX7 felis silve
24	2	28.6	8	8	Q94EX6	Q94EX6 felis libyc
25	2	28.6	8	9	Q8SBJ0	Q8SBJ0 bacterioph
26	2	28.6	8	9	Q8H9K1	Q8H9K1 bacterioph
27	2	28.6	8	10	Q8GKT5	Q8GKT5 lycopersico
28	2	28.6	8	11	Q60615	Q60615 mus musculu
29	2	28.6	8	13	Q902V5	Q902V5 fulica leuc
30	2	28.6	9	2	Q48687	Q48687 lactococcus
31	2	28.6	9	2	Q9RSM1	Q9RSM1 staphylococ
32	2	28.6	9	2	Q83157	Q83157 anabaena sp
33	2	28.6	9	2	Q83222	Q83222 streptomyce
34	2	28.6	9	2	Q8GL31	Q8GL31 borrelia bu
35	2	28.6	9	2	Q8GL26	Q8GL26 borrelia bu
36	2	28.6	9	3	Q9UR18	Q9UR18 sclerosium
37	2	28.6	9	4	Q96P97	Q96P97 homo sapien
38	2	28.6	9	6	Q9TRU7	Q9TRU7 bos taurus
39	2	28.6	9	6	Q28093	Q28093 bos taurus
40	2	28.6	9	6	Q9TRS0	Q9TRS0 oryctolagus
41	2	28.6	9	7	Q31415	Q31415 gallus gall
42	2	28.6	9	8	Q9T2L0	Q9T2L0 spinacia ol
43	2	28.6	9	8	Q8MBF4	Q8MBF4 ipomoea qua
44	2	28.6	9	8	Q94VG2	Q94VG2 varanus ind
45	2	28.6	9	8	Q9T2L1	Q9T2L1 spinacia ol
46	2	28.6	9	11	Q62530	Q62530 mus spretus
47	2	28.6	9	11	Q8CG13	Q8CG13 mus musculu
48	2	28.6	9	12	Q8QRR6	Q8QRR6 transmissib
49	2	28.6	9	12	Q8QRR5	Q8QRR5 transmissib
50	2	28.6	9	13	Q8AUM7	Q8AUM7 carassius a
51	2	28.6	10	2	Q9F9H5	Q9F9H5 helicobacte
52	2	28.6	10	2	Q9UN49	Q9UN49 staphylococ
53	2	28.6	10	2	Q9XBH3	Q9XBH3 bacillus ce
54	2	28.6	10	2	Q9ZIB1	Q9ZIB1 clostridium
55	2	28.6	10	2	Q8KIN9	Q8KIN9 clostridium
56	2	28.6	10	2	Q8VN85	Q8VN85 helicobacte
57	2	28.6	10	2	Q47475	Q47475 escherichia
58	2	28.6	10	2	Q54217	Q54217 staphylococ
59	2	28.6	10	2	Q9RSN2	Q9RSN2 clostridium
60	2	28.6	10	2	Q9RSN1	Q9RSN1 clostridium
61	2	28.6	10	2	Q9RSN3	Q9RSN3 clostridium
62	2	28.6	10	2	Q51812	Q51812 escherichia
63	2	28.6	10	2	Q83154	Q83154 anabaena sp
64	2	28.6	10	3	Q9UVW2	Q9UVW2 schizophyll
65	2	28.6	10	4	Q9UN90	Q9UN90 homo sapien
66	2	28.6	10	4	Q9PZ29	Q9PZ29 homo sapien
67	2	28.6	10	4	Q96QT9	Q96QT9 homo sapien
68	2	28.6	10	4	Q8WXB5	Q8WXB5 homo sapien
69	2	28.6	10	4	Q81ZA2	Q81ZA2 homo sapien
70	2	28.6	10	6	Q9TR47	Q9TR47 bos taurus
71	2	28.6	10	6	Q9TR48	Q9TR48 bos taurus
72	2	28.6	10	6	Q8HZ71	Q8HZ71 didelphis m
73	2	28.6	10	8	Q8MAZ9	Q8MAZ9 dicranostyl
74	2	28.6	10	8	Q8MBB7	Q8MBB7 merremia ae
75	2	28.6	10	10	P82937	P82937 hordeum vul
76	2	28.6	10	11	Q9QVK8	Q9QVK8 mus sp. mep
77	2	28.6	10	11	Q9R020	Q9R020 mus musculu
78	2	28.6	10	12	Q9Q0M9	Q9Q0M9 polyomaviru
79	2	28.6	10	12	Q8JV70	Q8JV70 polyomaviru
80	2	28.6	10	12	Q8JQW1	Q8JQW1 polyomaviru
81	2	28.6	10	12	Q8JV68	Q8JV68 polyomaviru
82	2	28.6	10	12	Q9Q0V9	Q9Q0V9 polyomaviru
83	2	28.6	10	12	Q90349	Q90349 hepatitis g
84	2	28.6	10	12	Q8JV84	Q8JV84 polyomaviru
85	2	28.6	10	12	Q9Q0W7	Q9Q0W7 polyomaviru
86	2	28.6	10	12	Q8JV66	Q8JV66 polyomaviru
87	2	28.6	10	12	Q9Q0V7	Q9Q0V7 polyomaviru
88	2	28.6	10	12	Q8JV82	Q8JV82 polyomaviru
89	2	28.6	10	12	Q8JV76	Q8JV76 polyomaviru

90 2 28.6 10 12 Q8JV74
 91 2 28.6 10 12 Q9Q0X7
 92 2 28.6 10 12 Q8JV78
 93 2 28.6 10 12 Q9Q0W5
 94 2 28.6 10 12 Q9Q0X3
 95 2 28.6 10 12 Q9Q0X5
 96 2 28.6 10 12 Q9Q0W3
 97 2 28.6 10 12 Q8JV80
 98 2 28.6 10 12 Q9Q0X1
 99 2 28.6 10 12 Q9Q0X9
 100 2 28.6 10 12 Q8JV72
 101 2 28.6 10 13 Q8PRY8
 102 2 28.6 10 15 Q85462
 103 1 14.3 4 11 Q08433
 104 1 14.3 5 2 P83073
 105 1 14.3 5 10 Q99007
 106 1 14.3 5 13 P83308
 107 1 14.3 6 10 P82181
 108 1 14.3 6 10 P82541
 109 1 14.3 6 10 P82182
 110 1 14.3 7 2 Q8KMS3
 111 1 14.3 7 2 P70804
 112 1 14.3 7 2 O50556
 113 1 14.3 7 2 Q47477
 114 1 14.3 7 2 Q47029
 115 1 14.3 7 2 P72081
 116 1 14.3 7 2 Q54248
 117 1 14.3 7 2 Q8KMS9
 118 1 14.3 7 2 Q34028
 119 1 14.3 7 2 Q8G112
 120 1 14.3 7 2 Q8GL04
 121 1 14.3 7 2 Q8GL00
 122 1 14.3 7 4 Q15903
 123 1 14.3 7 4 Q8NH7
 124 1 14.3 7 4 Q15897
 125 1 14.3 7 6 Q28742
 126 1 14.3 7 8 P92214
 127 1 14.3 7 8 P92393
 128 1 14.3 7 8 P92403
 129 1 14.3 7 8 P92427
 130 1 14.3 7 8 Q99182
 131 1 14.3 7 8 P92430
 132 1 14.3 7 8 P92221
 133 1 14.3 7 8 Q98866
 134 1 14.3 7 8 P92425
 135 1 14.3 7 8 P92381
 136 1 14.3 7 8 P92387
 137 1 14.3 7 8 P92210
 138 1 14.3 7 8 P92440
 139 1 14.3 7 8 P92218
 140 1 14.3 7 8 P92390
 141 1 14.3 7 8 P92372
 142 1 14.3 7 8 P92442
 143 1 14.3 7 8 P92226
 144 1 14.3 7 8 Q8MFV6
 145 1 14.3 7 8 Q95945
 146 1 14.3 7 8 P92385
 147 1 14.3 7 8 P92421
 148 1 14.3 7 10 P93233
 149 1 14.3 7 10 Q49223
 150 1 14.3 7 10 Q9C5B3

ALIGNMENTS

RESULT 1
 ID P70243 PRELIMINARY; PRT; 8 AA.
 AC P70243;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE Skeletal muscle-specific calcium channel (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129;
 RA Ophoff R.A.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: X98325; CAA66969.1; -.
 FT NON_TER 1 1
 FT NON_TER 8 8
 SQ SEQUENCE 8 AA; 865 MW; D9C37DDB1861ADDB CRC64;
 Query Match 42.9%; Score 3; DB 11; Length 8;
 Best Local Similarity 100.0%; Pred.No. 8.3e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TGE 3
 Db |||
 3 TGE 5
 RESULT 2
 ID Q62528 PRELIMINARY; PRT; 8 AA.
 AC Q62528;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
 DE Apolipoprotein A-II (APO-AII) (Fragment).
 GN APOA2
 OS Mus spretus (Western wild mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10096;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRET/EI;
 RX MEDLINE=94319082; PubMed=8043949;
 RA KO M.S., Wang X., Horton J.H., Takahashi N., Maesaki Y.,
 Nadeau J.H.;
 RL "Genetic mapping of 40 cDNA clones on the mouse genome by PCR.";
 Mamm. Genome 5:349-355(1994).
 CC -!- FUNCTION: MAY STABILIZE HDL (HIGH DENSITY LIPOPROTEIN) STRUCTURE
 BY ITS ASSOCIATION WITH LIPIDS, AND AFFECT THE HDL METABOLISM.
 DR EMBL: U05692; AAB60463.1; -.
 DR MGD; MGI:88050; APOA2.
 KW Plasma; Lipid transport; HDL.
 FT NON_TER 1 1
 FT NON_TER 8 8
 SQ SEQUENCE 8 AA; 827 MW; 223DDDD72DC7633B CRC64;
 Query Match 42.9%; Score 3; DB 11; Length 8;
 Best Local Similarity 100.0%; Pred.No. 8.3e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 EKP 5
 Db |||
 1 EKP 3
 RESULT 3
 ID Q9EZ14 PRELIMINARY; PRT; 9 AA.
 AC Q9EZ14;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE SPAS (Fragment).
 GN SPAS.
 OS Sodalis glossinidius.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

```

OC Enterobacteriaceae; Sodalis.
OX NCBI_TaxID=63612;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M1;
RX MEDLINE=21117132; PubMed=11172045;
RA Dale C., Young S.A., Haydon D.T., Welburn S.C.;
RT "The insect endosymbiont Sodalis glossinidius utilizes a type III
RT secretion system for cell invasion.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:1883-1888 (2001).
DR EMBL; AF306650; AAG48607.1; -.
FT NON TER 9
SQ SEQUENCE 9 AA; 1035 MW; 818E633B01A33455 CRC64;

Query Match 42.9%; Score 3; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EKP 5
Db [1]
6 EKP 8

RESULT 4
Q9R735 PRELIMINARY; PRT; 9 AA.
AC Q9R735;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Fkba protein (Fragment).
GN Fkba.
OS Streptomyces chrysomallus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1899;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94341259; PubMed=8062824;
RA Pahl A., Keller U.;
RT "Streptomyces chrysomallus FKBP-33 is a novel immunophilin consisting
RT of two FKBP-binding domains with its gene transcriptionally coupled
RT to the FKBP-12 gene.";
RL EMBO J. 13:3472-3480 (1994).
DR EMBL; Z34523; CAA84282.1; -.
FT NON TER 9
SQ SEQUENCE 9 AA; 1047 MW; 9A8BCB07633B1045 CRC64;

Query Match 42.9%; Score 3; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EKP 5
Db [1]
4 EKP 6

RESULT 5
Q16605 PRELIMINARY; PRT; 9 AA.
AC Q16605;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Glutathione S-transferase 2 (Fragment).
GN GSTA2 OR GST2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88330756; PubMed=3138230;

Chow N.W., Whang-Peng J., Kao-Shan C.S., Tam M.F., Lai H.C., Tu C.P.;
RA "Human glutathione S-transferases. The Ha multigene family encodes
RT products of different but overlapping substrate specificities.";
RL J. Biol. Chem. 263:12797-12800 (1988).
DR EMBL; M21867; AAA52617.1; -.
DR EMBL; M21866; AAA35938.1; -.
FT NON TER 9
SQ SEQUENCE 9 AA; 1116 MW; D168E7232763381D CRC64;

Query Match 42.9%; Score 3; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EKP 5
Db [1]
3 EKP 5

RESULT 6
Q9QZA8 PRELIMINARY; PRT; 9 AA.
AC Q9QZA8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE C-type lectin DCL1 (Fragment).
GN DCL1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX Gorski K., Huang X., Tseng S.-Y., Rattis F., Pardoll D., Teuchiya H.;
RT "Dendritic cell regulation of DCL1 mRNA expression.";
RL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF192526; AAF04843.1; -.
DR MGD; MGI:2136650; DCL1.
KW Lectin.
FT NON TER 9
SQ SEQUENCE 9 AA; 994 MW; 342161AB172BAB7 CRC64;

Query Match 42.9%; Score 3; DB 11; Length 9;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
Db [1]
7 TGE 9

RESULT 7
Q47505 PRELIMINARY; PRT; 7 AA.
AC Q47505;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MCCA protein.
GN MCCA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9609297; PubMed=8522520;
RA Gonzalez-Pascor J.E., San Millan J.L., Castilla M.A., Moreno F.;
RT "Structure and organization of plasmid genes required to produce the
RT translation inhibitor microcin C7.";
RL J. Bacteriol. 177:7131-7140 (1995).
DR EMBL; X57583; CAA40808.1; -.

```

KW Plasmid.
SQ SEQUENCE 7 AA; 763 MW; 644DD44861B406F0 CRC64;
Query Match 28.6%; Score 2; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TG 2
||
Db 3 TG 4
||
RESULT 8
Q66205 PRELIMINARY; PRT; 7 AA.
AC Q66205;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Membrane protein (1 is 3rd base in codon) (Fragment).
OS Transmissible gastroenteritis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OC NCBI_TaxID=11149;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FS772/70;
RX MEDLINE=86216185; PubMed=2835592;
RA Britton P., Carnes R.S., Page K.W., Garves D.J., Parra F.;
RT "Sequence of the Nucleoprotein Gene from a Virulent British Field
RT Isolate of Transmissible Gastroenteritis Virus and its Expression in
RT Saccharomyces Cerevisiae.";
RL Mol. Microbiol. 2:89-99(1988).
DR EMBL; Y00542; CAA68606.1; -.
FT NON_TER 1
SQ SEQUENCE 7 AA; 869 MW; 72C691E772733B70 CRC64;
Query Match 28.6%; Score 2; DB 12; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 EK 4
||
Db 1 EK 2
||
RESULT 9
Q9R772 PRELIMINARY; PRT; 8 AA.
AC Q9R772;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 1.0 kDa protein (Fragment).
GN YOFG.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayaishi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA Sequence of Escherichia coli K-12 Genome Corresponding
RT to the 12.7-28.0 min Region on the Linkage Map.";
RL DNA Res. 3:137-155(1996).
DR EMBL; D90705; BAA35310.1; -.

KW Hypothetical protein.
SQ SEQUENCE 8 AA; 964 MW; DFL33B1DD04B476A CRC64;
Query Match 28.6%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 EK 4
||
Db 6 EK 7
||
RESULT 10
Q9AGP4 PRELIMINARY; PRT; 8 AA.
AC Q9AGP4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Serine hydroxymethyltransferase (Fragment).
GN GUYA.
OS Arthrobacter SP. 11N.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococineae; Micrococcaceae; Arthrobacter.
OX NCBI_TaxID=153502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=11N;
RA Meskys R., Harris R.J., Casalte V., Bastian J., Scrutton N.S.;
RT "Genetic organization of the genes involved in dimethylglycine and
RT sarcosine degradation in Arthrobacter spp.: implications for glycine
RT betaine catabolism.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329478; AAK16486.1; -.
FT NON_TER 1
SQ SEQUENCE 8 AA; 898 MW; 6B18705333372457 CRC64;
Query Match 28.6%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 GE 3
||
Db 7 GE 8
||
RESULT 11
Q44463 PRELIMINARY; PRT; 8 AA.
AC Q44463;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Agrobacterium tumefaciens Ti plasmid virD3 and virD4 genes
DE (Fragment).
OS Rhizobiales (rhizobacteria).
OX Bacteria; Proteobacteria; Alphaproteobacteria.
OX NCBI_TaxID=356;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88015611; PubMed=3658701;
RA Portner S.G., Yanofsky M.F., Nester E.W.;
RT "Molecular characterization of the virD operon from Agrobacterium
RT tumefaciens.";
RL Nucleic Acids Res. 15:7503-7517(1987).
DR EMBL; X6045; CAA29439.1; -.
FT NON_TER 8
SQ SEQUENCE 8 AA; 887 MW; F8F2C325B33861A6 CRC64;

Query Match 28.6%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
||
DB 2 TG 3

RESULT 12

ID Q16428 PRELIMINARY; PRT; 8 AA.
AC Q16428;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Dystrophin protein (fragment).
GN DYSTROPHIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96163501; PubMed=8566960;
RA Holder E., Maeda M., Bies R.D.;
RT "Expression and regulation of the dystrophin Purkinje promoter in
RT human skeletal muscle, heart, and brain."
RL Hum. Genet. 97:232-239 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Kitano T., Kobayakawa H., Saitou N.;
RL "Silver Project."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; S81419; AAD14363.1; --
DR EMBL; AB037493; BAA50413.1; --
FT NON_TER 8
SQ SEQUENCE 8 AA; 825 MW; F21865A5A2CB05B6 CRC64;

Query Match 28.6%; Score 2; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GE 3
||
DB 7 GE 8

RESULT 13

ID Q15888 PRELIMINARY; PRT; 8 AA.
AC Q15888;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE (Clone XP1588A) (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISUES=Placenta;
RA Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA Coolbaugh M.I., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,
RA Caskey C.T.H.;
RT "Isolation of chromosome-specific genes by reciprocal probing of
RT arrayed cDNAs and cosmid libraries."
RL Hum. Mol. Genet. 0:0-0 (1995).
DR EMBL; L32069; AAA73878.1; --
FT NON_TER 1
FT NON_TER 8
SQ SEQUENCE 8 AA; 1068 MW; 0315A337EAB5B0763 CRC64;

Query Match 28.6%; Score 2; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KP 5
||
DB 1 KP 2

RESULT 14

ID Q9UJ50 PRELIMINARY; PRT; 8 AA.
AC Q9UJ50;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Latrophilin-2 (fragment).
GN LPPH1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99153747; PubMed=10030676;
RA White G.R.M., Varley J.M., Heighway J.;
RT "Isolation and characterisation of a human homologue of the
RT latrophilin gene from a region of lp31.1 implicated in breast
RT cancer."
RL Oncogene 17:3513-3519 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20225451; PubMed=10760572;
RA White G.R.M., Varley J.M., Heighway J.;
RT "Genomic structure and expression profile of LPPH1, a 7TM gene
RT variably expressed in breast cancer cell lines."
RL Biochim. Biophys. Acta 1491:75-92 (2000).
DR EMBL; AJ244509; CAB60204.1; --
FT NON_TER 1
FT NON_TER 8
SQ SEQUENCE 8 AA; 1025 MW; 4D59D763344AAB0B CRC64;

Query Match 28.6%; Score 2; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KP 5
||
DB 5 KP 6

RESULT 15

ID Q9TWH6 PRELIMINARY; PRT; 8 AA.
AC Q9TWH6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE BIOACTIVE peptide P4=PUTATIVE ESOPHAGEAL NEUROREGULATOR.
OS Perinereis vancaurica.
OC Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Aciculata;
OC Phylodocida; Nereididae; Perinereis.
OX NCBI_TaxID=6355;
RN [1]
RP SEQUENCE
RX MEDLINE=95323338; PubMed=7599979;
RA Takahashi T., Furukawa Y., Muneoka Y., Matsushima O., Ikeda T.,
RA Fujita T., Minakata H., Nomoto K.;
RT "Isolation and characterization of four novel bioactive peptides from
RT a polychaete annelid, Perinereis vancaurica."
RL Comp. Biochem. Physiol. C, Pharmacol. Toxicol. Endocrinol. 110:297-304 (1995).

SQ SEQUENCE 8 AA; 989 MW; 954772CAA87B0B59 CRC64;

Query Match 28.6%; Score 2; DB 5; Length 8;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PY 6
 ||
 Db 7 PY 8

RESULT 16

Q94695 PRELIMINARY; PRT; 8 AA.

AC Q94695;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
 DE Actin (Fragment).
 GN ARDC.
 OS Physarum polycephalum (Slime mold).
 OC Eukaryota; Mycetozoa; Myxogastria; Myxogastromycetidae; Physarida;
 OC Physarum.
 OX NCBI_TaxID=5791;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96182101; PubMed=8622700;
 RA Benard M., Lagnel C., Pallotta D., Pierron G.;
 RT "Mapping of a replication origin within the promoter region of two
 RT unlinked, abundantly transcribed actin genes of Physarum
 RT polycephalum.";
 RL Mol. Cell. Biol. 16:968-976(1996).
 DR EMBL: M73459; AAB03706.1; -;
 FT NON TER 8
 SQ SEQUENCE 8 AA; 878 MW; F4C6C2CAAB187B16 CRC64;

Query Match 28.6%; Score 2; DB 5; Length 8;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GE 3
 ||
 Db 3 GE 4

RESULT 17

Q9TRX8 PRELIMINARY; PRT; 8 AA.

AC Q9TRX8;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Osteopontin (Fragment).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=91282766; PubMed=1676261;
 RA Prince C.W., Dickie D., Krundieck C.L.;
 RT "Osteopontin, a substrate for transglutaminase and factor XIII
 RT activity.";
 RL Biochem. Biophys. Res. Commun. 177:1205-1210(1991).
 FT NON TER 1 1
 FT NON TER 8 8
 SQ SEQUENCE 8 AA; 920 MW; 05DAFAF76322D767 CRC64;

Query Match 28.6%; Score 2; DB 6; Length 8;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KP 5
 ||
 Db 4 KP 5

RESULT 18

Q9XSY1 PRELIMINARY; PRT; 8 AA.

AC Q9XSY1;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
 DE Retinoblastoma protein (Fragment).
 GN Rbl.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97049323; PubMed=8894053;
 RA Venter P.J., Brouillette J.A., Yuzbasiyan-Gurkan V., Brewer G.J.;
 RT "Gene-specific universal mammalian sequence-tagged sites: application
 RT to the canine genome.";
 RL Biochem. Genet. 34:321-341(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Venter P.J., Cao Y., Alexander L., Yuzbasiyan-Gurkan V.;
 RT "Dinucleotide repeat polymorphism in the canine retinoblastoma (Rb1)
 RT gene.";
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBSJ databases.
 DR EMBL: AF155737; AAD38807.1; -;
 FT NON TER 1 1
 FT NON TER 8 8
 SQ SEQUENCE 8 AA; 895 MW; 1425BB18676721E3 CRC64;

Query Match 28.6%; Score 2; DB 6; Length 8;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GE 3
 ||
 Db 5 GE 6

RESULT 19

Q9MYL5 PRELIMINARY; PRT; 8 AA.

AC Q9MYL5;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE Dystrophin.
 OS Pongo pygmaeus (Orangutan).
 OS Pan troglodytes (Chimpanzee), and
 OS Gorilla gorilla (gorilla).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pongo.
 OX NCBI_TaxID=9600, 9598, 9593;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=P. pygmaeus, P. troglodytes, and G. gorilla;
 RA Kitano T., Kobayakawa H., Saitou N.;
 RT "Silver Project.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBSJ databases.
 DR EMBL: AB037496; BAA90419.1; -;
 DR EMBL: AB037494; BAA90415.1; -;
 DR EMBL: AB037495; BAA90417.1; -;
 SQ SEQUENCE 8 AA; 825 MW; F21865A5A2CB05B6 CRC64;

Query Match 28.6%; Score 2; DB 6; Length 8;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GE 3
DB 7 GE 8

RESULT 20

Q95M23 PRELIMINARY; PRT; 8 AA.

AC Q95M23;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Na+/K+ ATPase alpha subunit (EC 3.6.1.37) (Fragment).
GN ATP1A1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pietrain;
RA Blazkova P., Stratil A., Peelman L.J., Van Poucke M., Reiner G.,
RA Geldermann H., Kopecky M.;
RT "RH mapping of the porcine ATP1A1, ATP1B1, V-ATPase, IVL genes and
RT linkage assignments of ATP1A1 and IVL to chromosome 4.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ344138; CAC51422.1; -;
DR EMBL; AJ344138; CAC51422.1; -;
KW Hydrolase.
FT NON_TER 1
SQ SEQUENCE 8 AA; 1117 MW; 604B41AB133B02D3 CRC64;

Query Match 28.6%; Score 2; DB 6; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EK 4
DB 3 EK 4

RESULT 21

Q95213 PRELIMINARY; PRT; 8 AA.

AC Q95213;
DT 01-FEB-1997 (TReMBLrel. 02, Created)
DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Germline DH (DF) gene (Fragment).
GN DF.
OC Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R-1/rgm;
RA Mage R.G., Chen H.-T., Alexander C.B., Chen F.F.;
RT "Rabbit DQ52 and DH Gene Rearrangements in Early B-cell Development.";
RL Mol. Immunol. 0:0-0(1996).
DR EMBL; U62585; AAB18735.1; -;
FT NON_TER 1
FT NON_TER 8
SQ SEQUENCE 8 AA; 845 MW; 5CA861B5AB58677B CRC64;

Query Match 28.6%; Score 2; DB 7; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
DB 6 TG 7

RESULT 22

Q94PX5 PRELIMINARY; PRT; 8 AA.

AC Q94PX5;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Cytochrome c oxidase subunit II (Fragment).
GN COII.
OS Felis silvestris (Wild cat).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9683;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=66, 71, 75, 90, 1, and 2;
RA Randi E., Pierpaoli M., Beaumont M., Ragni B., Sforzi A.;
RT "Genetic identification of wild and domestic cats (Felis silvestris),
RT and their hybrids using Bayesian clustering methods.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ409136; CAC41051.1; -;
DR EMBL; AJ409137; CAC41054.1; -;
DR EMBL; AJ409138; CAC41057.1; -;
DR EMBL; AJ409139; CAC41060.1; -;
DR EMBL; AJ409141; CAC41066.1; -;
DR EMBL; AJ409143; CAC41072.1; -;
KW Mitochondrion.
FT NON_TER 1
SQ SEQUENCE 8 AA; 951 MW; 262685BDC5A3733B CRC64;

Query Match 28.6%; Score 2; DB 8; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EK 4
DB 1 EK 2

RESULT 23

Q94PX7 PRELIMINARY; PRT; 8 AA.

AC Q94PX7;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Cytochrome c oxidase subunit II (Fragment).
GN COII.
OS Felis silvestris catus (Cat).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1, 2, 7, 12, 16, 17, and 110;
RA Randi E., Pierpaoli M., Beaumont M., Ragni B., Sforzi A.;
RT "Genetic identification of wild and domestic cats (Felis silvestris),
RT and their hybrids using Bayesian clustering methods.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ409128; CAC41027.1; -;
DR EMBL; AJ409129; CAC41030.1; -;
DR EMBL; AJ409130; CAC41033.1; -;
DR EMBL; AJ409131; CAC41036.1; -;
DR EMBL; AJ409132; CAC41039.1; -;
DR EMBL; AJ409133; CAC41042.1; -;
DR EMBL; AJ409134; CAC41045.1; -;
KW Mitochondrion.
FT NON_TER 1
SQ SEQUENCE 8 AA; 951 MW; 262685BDC5A3733B CRC64;

Query Match 28.6%; Score 2; DB 8; Length 8;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EK 4
 ||
 Db 1 EK 2

RESULT 24
 Q94PX6 PRELIMINARY; PRT; 8 AA.
 AC Q94PX6; DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Cytochrome c oxidase subunit II (Fragment).
 GN COII; libyca.
 OS Felis libyca.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
 OX NCBI_TaxID=61377;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=40, 1, 2, and 7;
 RA Randi E., Pierpaoli M., Beaumont M., Ragni B., Sforzi A.;
 RT "Genetic identification of wild and domestic cats (*Felis silvestris*),
 and their hybrids using Bayesian clustering methods."
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ409135; CAC41048.1; -
 DR EMBL; AJ409140; CAC41063.1; -
 DR EMBL; AJ409142; CAC41069.1; -
 DR EMBL; AJ409144; CAC41075.1; -
 KW Mitochondrion.
 FT NON_TER 1
 SQ SEQUENCE 8 AA; 951 MW; 262685BDC5A3733B CRC64;

Query Match 28.6%; Score 2; DB 8; Length 8;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EK 4
 ||
 Db 1 EK 2

RESULT 25
 Q8SBJ0 PRELIMINARY; PRT; 8 AA.
 AC Q8SBJ0; DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Gp30.2 (Fragment).
 GN Gp30.2.
 OS Bacteriophage RB69.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
 OC T4-like viruses.
 OX NCBI_TaxID=12353;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Piesiniene L., Kolesinskiene G., Truncaite L., Zajackauskaite A.,
 RA Nivinskas R.;
 RT "Genomic region with genes 30.6-30.3 of T4-related bacteriophages."
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ439452; CAD28423.1; -
 FT NON_TER 8
 SQ SEQUENCE 8 AA; 918 MW; F3B1A72041B76336 CRC64;

Query Match 28.6%; Score 2; DB 9; Length 8;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KP 5
 ||
 Db 2 KP 3

RESULT 26
 Q8H9K1 PRELIMINARY; PRT; 8 AA.
 AC Q8H9K1; DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Gp30.2 protein (Fragment).
 GN Gp30.2.
 OS Bacteriophage L210.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
 OC T4-like viruses.
 OX NCBI_TaxID=192973;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kolesinskiene G., Nivinskas R.;
 RT "A pair of overlapping genes 30.3 and 30.3' of T4-related
 bacteriophages."
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ458400; CAD30256.1; -
 FT NON_TER 8
 SQ SEQUENCE 8 AA; 918 MW; F3B1A72041B76336 CRC64;

Query Match 28.6%; Score 2; DB 9; Length 8;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KP 5
 ||
 Db 2 KP 3

RESULT 27
 Q8GTG5 PRELIMINARY; PRT; 8 AA.
 AC Q8GTG5; DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE WRKY transcription factor IID-1 splice variant 2 (Fragment).
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; lamids; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4081;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Money Maker; TISSUE=Leaf;
 RA Cagna G., Boes A., Ulker B., Zhou A., Wanke D., Somssich I.E.;
 RT "Comparison of WRKY group II transcription factors from plants."
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY157059; AAN71729.1; -
 FT NON_TER 1
 SQ SEQUENCE 8 AA; 917 MW; 5917B4775B87330 CRC64;

Query Match 28.6%; Score 2; DB 10; Length 8;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PY 6
 ||
 Db 5 PY 6

RESULT 28
 Q60615 PRELIMINARY; PRT; 8 AA.
 ID Q60615

AC O0615;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
 DE Apolipoprotein A-II (Fragment).
 GN APOA2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=94319082; PubMed=8043949;
 RA Ko M.S., Wang X., Horton J.H., Hagen M.D., Takahashi N., Maerzaki Y.,
 RA Nadeau J.H.;
 RT "Genetic mapping of 40 cDNA clones on the mouse genome by PCR."
 RL Mamm. Genome 5:349-355(1994).
 DR EMBL; U05691; AAB60462.1; -;
 DR MGD; MGI:88050; APOA2.
 KW Lipoprotein.
 FT NON_TER 1
 SQ SEQUENCE 8 AA; 720 MW; 783DDC5B861AB18 CRC64;

Query Match 28.6%; Score 2; DB 11; Length 8;
 Best Local Similarity 100.0%; Pred.No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GE 3
 ||
 Db 1 GE 2

RESULT 29

ID Q90ZV5 PRELIMINARY; PRT; 8 AA.
 AC Q90ZV5;
 DT 01-DEC-2001 (TRENBLrel. 19, Created)
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE Adenylate kinase (Fragment).
 OS Pulica leucoptera.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Gruiformes; Rallidae; Fulica.
 OX NCBI_TaxID=156758;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Shapero L.H., Dumbacher J.P.;
 RT "Adenylate kinase intron 5: a new nuclear locus for avian
 RT systematics."
 RL Auk 118:248-255(2001).
 DR EMBL; AF307898; AAK43537.1; -;
 KW Kinase.
 FT NON_TER 1
 FT NON_TER 8
 SQ SEQUENCE 8 AA; 994 MW; 96333B19CB1B1866 CRC64;

Query Match 28.6%; Score 2; DB 13; Length 8;
 Best Local Similarity 100.0%; Pred.No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GE 3
 ||
 Db 2 GE 3

RESULT 30

ID Q48687 PRELIMINARY; PRT; 9 AA.
 AC Q48687;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)

DE Streptococcus cremoris promoter 32 DNA (Fragment).
 OS Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 OX NCBI_TaxID=1359;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88105390; PubMed=2447829;
 RA van der Vossen J.M.B.M., van der Leije D., Venema G.;
 RT "Isolation and characterization of Streptococcus cremoris WG2-specific
 RT promoters."
 RL Appl. Environ. Microbiol. 53:2452-2457(1987).
 DR EMBL; M24764; AAA26961.1; -;
 FT NON_TER 9
 SQ SEQUENCE 9 AA; 995 MW; C3FA3B1DC5A2C05D CRC64;

Query Match 28.6%; Score 2; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred.No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EK 4
 ||
 Db 7 EK 8

RESULT 31

ID Q9R5M1 PRELIMINARY; PRT; 9 AA.
 AC Q9R5M1;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE 66 kDa cell surface adhesin for heparan sulfate (Fragment).
 OS Staphylococcus aureus.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=1280;
 RN [1]
 RP SEQUENCE.
 RA Liang O.D., Ascencio F., Fransson L.A., Wadstrom T.;
 RT "Binding of heparan sulfate to Staphylococcus aureus."
 RL Infect. Immun. 60:899-906(1992).
 FT NON_TER 1
 FT NON_TER 9
 SQ SEQUENCE 9 AA; 990 MW; 2289DD7337861B3 CRC64;

Query Match 28.6%; Score 2; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred.No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
 ||
 Db 3 TG 4

RESULT 32

ID P83157 PRELIMINARY; PRT; 9 AA.
 AC P83157;
 DT 01-DEC-2001 (TRENBLrel. 19, Created)
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE Superoxide dismutase [Fe] (EC 1.15.1.1) (Fragment).
 OS Anabaena sp. (strain L31).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.
 OX NCBI_TaxID=29412;
 RN [1]
 RP SEQUENCE.

RA Apte S.K., Uhlemann E., Schmid R., Altendorf K.;
 RL Submitted (OCT-2001) to the SWISS-PROT data bank.
 CC -!- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
 CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.
 CC -!- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) = O(2) + H(2)O(2).
 CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -!- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
 CC FAMILY.
 DR InterPro: IPR001189; SODismutase.
 DR PROSITE: PS00088; SOD MN: PARTIAL.
 KW Oxidoreductase; Iron; Metal-binding.
 FT NON_TER 9
 SQ SEQUENCE 9 AA; 1063 MW; C54267376B06C2C9 CRC64;

Query Match 28.6%; Score 2; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YK 6
 DB 8 YK 9

RESULT 33
 P83222 PRELIMINARY; PRT; 9 AA.
 ID P83222;
 AC P83222;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE 34 kDa extracellular nuclease (EC 3.-.-.-) (Fragment).
 OS Streptomyces antibioticus.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxID=1890;
 RN [1]
 RP SEQUENCE, FUNCTION, ENZYME REGULATION, SUBCELLULAR LOCATION, AND
 RP DEVELOPMENTAL STAGE.
 RC STRAIN=ATCC 11891;
 RX MEDLINE=99329048; PubMed=10400660;
 RA Nicieza R.C., Huergo J., Connolly B.A., Sanchez J.;
 RT "Purification, characterization, and role of nucleases and serine
 RT proteases in Streptomycetes differentiation. Analogies with the
 RT biochemical processes described in late steps of eukaryotic
 RT apoptosis";
 RL J. Biol. Chem. 274:20366-20375 (1999).
 CC -!- FUNCTION: INVOLVED IN DNA DEGRADATION IN THE SUBSTRATE MYCELIUM.
 CC CUTS DNA NONSPECIFICALLY. POSSESSES ENONUCLEOLYTIC ACTIVITY.
 CC -!- ENZYME REGULATION: STIMULATED BY MAGNESIUM AND MANGANESE.
 CC INHIBITED BY ZINC AND AURIN TRICARBOXYLIC ACID.
 CC -!- SUBUNIT: MONOMER (PROBABLE).
 CC -!- SUBCELLULAR LOCATION: SECRETED.
 CC -!- DEVELOPMENTAL STAGE: HIGHEST EXPRESSION FOUND DURING AERIAL
 CC MYCELIUM FORMATION AND SPOULATION.
 CC -!- MISCELLANEOUS: OPTIMUM PH IS 8-8.5.
 KW Hydrolase; Nuclease; Magnesium.
 FT UNSURE 1 1 OR E.
 FT UNSURE 5 5 OR A.
 FT NON_TER 9 9
 SQ SEQUENCE 9 AA; 994 MW; 80B05AB6D8705731 CRC64;

Query Match 28.6%; Score 2; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 YK 7
 DB 8 YK 9

RESULT 34
 Q8GL31 PRELIMINARY; PRT; 9 AA.
 ID Q8GL31;
 AC Q8GL31;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE PF-50 protein (Fragment).

GN PF-50.
 OS Borrelia burgdorferi (lyme disease spirochete).
 OG Plasmid group cp32-1.
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sh-2-82;
 RA Stevenson B., Miller J.C.;
 RT "Comparative analyses of Borrelia burgdorferi erp genes and their cp32
 RT prophages: conservation amidst diversity";
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY142089; AAN17869.1; -.
 KW Plasmid.
 FT NON_TER 1 1
 SQ SEQUENCE 9 AA; 1206 MW; 5A4A244337204373 CRC64;

Query Match 28.6%; Score 2; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 YK 7
 DB 1 YK 2

RESULT 35
 Q8GL26 PRELIMINARY; PRT; 9 AA.
 ID Q8GL26;
 AC Q8GL26;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE PF-50 protein (Fragment).
 GN PF-50.
 OS Borrelia burgdorferi (lyme disease spirochete).
 OG Plasmid group cp32-5.
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sh-2-82;
 RA Stevenson B., Miller J.C.;
 RT "Comparative analyses of Borrelia burgdorferi erp genes and their cp32
 RT prophages: conservation amidst diversity";
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY142092; AAN17873.1; -.
 KW Plasmid.
 FT NON_TER 1 1
 SQ SEQUENCE 9 AA; 1206 MW; 5A4A244330504373 CRC64;

Query Match 28.6%; Score 2; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 YK 7
 DB 1 YK 2

RESULT 36
 Q9UR18 PRELIMINARY; PRT; 9 AA.
 ID Q9UR18;
 AC Q9UR18;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE Translation elongation factor 1-alpha (Fragment).
 GN EFL-ALPHA.
 OS Sclerotium cepivorum.
 OC Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Sclerotium.
 OX NCBI_TaxID=38492;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=DE-2, and DE-1;
 RA Couch B.C., Kohn L.M.;
 RT "Clonal spread of Sclerotium cepivorum in onion production with
 evidence of past recombination events."
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR ENBL: AF163671; AAD46912.1; --
 DR ENBL: AF163670; AAD46911.1; --
 KW Elongation factor.
 FT NON TER 1 1
 FT NON TER 9 9
 SQ SEQUENCE 9 AA; 1178 MW; 619C19C3381041B4 CRC64;

Query Match 28.6%; Score 2; DB 3; Length 9;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EK 4
 ||
 Db 5 EK 6

RESULT 37

Q96P97 PRELIMINARY; PRT; 9 AA.
 AC Q96P97;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Reptin52 protein (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kang H.S., Park Y.J., Jung H.M., Jun D.Y., Huh T.L., Kim Y.H.;
 RT "Characterization of TPA-responsive genes in U937 cells using ordered
 RT differential display PCR."
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR ENBL: AF401216; AAL02172.1; --
 FT NON TER 1 1
 FT NON TER 9 AA; 981 MW; 5CDDAA681AB1873 CRC64;

Query Match 28.6%; Score 2; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GE 3
 ||
 Db 3 GE 4

RESULT 38

Q9TRU7 PRELIMINARY; PRT; 9 AA.
 AC Q9TRU7;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE GAP-3, GTPase-activating protein (Fragment).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE
 RX MEDLINE=92112868; PubMed=1309786;
 RA Nice E.C., Fabri L., Hammacher A., Holden J., Simpson R.J.,
 RA Burgess A.W.;
 RT "The purification of a Rap1 GTPase-activating protein from bovine
 brain cytosol.";

RL J. Biol. Chem. 267:1546-1553 (1992).
 FT NON TER 1 1
 FT NON TER 9 9
 SQ SEQUENCE 9 AA; 1063 MW; 89EDA77B47604B5A CRC64;
 Query Match 28.6%; Score 2; DB 6; Length 9;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PY 6
 ||
 Db 5 PY 6

RESULT 39

Q28093 PRELIMINARY; PRT; 9 AA.
 AC Q28093;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Cyokeratin IV gene upstream region (Fragment).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89231609; PubMed=2469572;
 RA Blessing M., Jorcano J.L., Franke W.W.;
 RT "Enhancer elements directing cell-type-specific expression of
 RT cyokeratin genes and changes of the epithelial cytoskeleton by
 RT transfections of hybrid cyokeratin genes."
 RL EMBO J. 8:117-126 (1989).
 DR ENBL: X14478; CAA32640.1; --
 KW Keratin.
 FT NON TER 9 9
 FT NON TER 9 AA; 1044 MW; 819A22D1B5B32B45 CRC64;

Query Match 28.6%; Score 2; DB 6; Length 9;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 YK 7
 ||
 Db 3 YK 4

RESULT 40

Q9TRSO PRELIMINARY; PRT; 9 AA.
 AC Q9TRSO;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Calcyclin-associated protein, CAP50=CA2+/phospholipid-binding protein
 L-7 fragment (Fragment).
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=92250478; PubMed=1533622;
 RA Tokumitsu H., Mizutani A., Minami H., Kobayashi R., Hidaka H.;
 RT "A calcyclin-associated protein is a newly identified member of the
 RT Ca2+/phospholipid-binding proteins, annexin family.";
 RL J. Biol. Chem. 267:8919-8924 (1992).
 FT NON TER 1 1
 FT NON TER 9 9
 SQ SEQUENCE 9 AA; 1010 MW; 64E419C44865B72B CRC64;

Query Match 28.6%; Score 2; DB 6; Length 9;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EK 4
 ||
 Db 8 EK 9

RESULT 41
 Q31415 PRELIMINARY; PRT; 9 AA.
 ID Q31415
 AC Q31415
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-JAN-1999 (TREMBlrel. 09, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE MHC class I antigen (Fragment).
 OS Gallus gallus (Chicken)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kanki T., Kuwasawa N., Sekiya Y., Ichikawa Y.;
 RT "Responsive expression of a MHC class I epitope and genes following
 RT Marek's disease virus infection."
 RL Submitted (JAN-1992) to the EMBL/GenBank/DBJ databases.
 DR EMBL; D90399; BAA14395.1; -;
 FT NON_TER 1
 SQ SEQUENCE 9 AA; 859 MW; 8A55A76455B861B5 CRC64;

Query Match 28.6%; Score 2; DB 7; Length 9;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
 ||
 Db 3 TG 4

RESULT 42
 Q9T2L0 PRELIMINARY; PRT; 9 AA.
 ID Q9T2L0
 AC Q9T2L0
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE LHC II-light-harvesting chlorophyll protein II (Fragment).
 OS Spinacia oleracea (Spinach).
 OC Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
 OX NCBI_TaxID=3562;
 RN [1]
 RP SEQUENCE.
 RA MEDLINE=91373386; PubMed=1894641;
 RA Michel H., Griffin P.R., Shabanowitz J., Hunt D.F., Bennett J.;
 RT "Tandem mass spectrometry identifies sites of three post-translational
 RT modifications of spinach light-harvesting chlorophyll protein II.
 RT Proteolytic cleavage, acetylation, and phosphorylation."
 RL J. Biol. Chem. 266:17584-17591(1991).
 FT NON_TER 1
 FT NON_TER 9
 SQ SEQUENCE 9 AA; 999 MW; 54D9276387DC1A3 CRC64;

Query Match 28.6%; Score 2; DB 8; Length 9;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KP 5
 ||

Db 6 KP 7
 RESULT 43
 Q8MBF4 PRELIMINARY; PRT; 9 AA.
 ID Q8MBF4
 AC Q8MBF4
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE P8BJ (Fragment).
 OS PSBJ.
 OS Ipomoea quamoclit.
 OC Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; Lamiales; Solanales; Convolvulaceae; Ipomoea.
 OX NCBI_TaxID=89660;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Stefanovic S., Krueger L., Olmstead R.G.;
 RT "Monophyly of the Convolvulaceae and circumscription of their major
 RT lineages based on DNA sequences of multiple chloroplast loci."
 RL Am. J. Bot. 0:0-0(2002).
 DR EMBL; AY100854; AAM55540.1; -;
 KW Chloroplast.
 FT NON_TER 9
 SQ SEQUENCE 9 AA; 961 MW; DD59440861B1AAAD CRC64;

Query Match 28.6%; Score 2; DB 8; Length 9;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
 ||
 Db 5 TG 6

RESULT 44
 Q94VG2 PRELIMINARY; PRT; 9 AA.
 ID Q94VG2
 AC Q94VG2
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Cytochrome c oxidase subunit I (Fragment).
 GN COI.
 OS Varanus indicus.
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
 OX NCBI_TaxID=62043;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ast J.C.;
 RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
 RL Cladistics 17:0-0(2001).
 DR EMBL; AF407505; AAL10069.1; -;
 KW Mitochondrion.
 FT NON_TER 9
 SQ SEQUENCE 9 AA; 1258 MW; 8B1259C727336411 CRC64;

Query Match 28.6%; Score 2; DB 8; Length 9;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 YK 7
 ||
 Db 8 YK 9

RESULT 45
 Q9T2L1

ID Q9T2L1 PRELIMINARY; PRT; 9 AA.
 AC Q9T2L1;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE LHC II=LIGHT-harvesting chlorophyll protein II (Fragment).
 OS Spinacia oleracea (Spinach).
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
 OX NCBI_TaxID=3562;
 RN [1]
 RP SEQUENCE
 RX MEDLINE=91373386; PubMed=1894641;
 RA Michel H., Griffin P.R., Shabanowitz J., Hunt D.F., Bennett J.;
 RT "Tandem mass spectrometry identifies sites of three post-translational
 RT modifications of spinach light-harvesting chlorophyll protein II.
 RT Proteolytic cleavage, acetylation, and phosphorylation.";
 RL J. Biol. Chem. 266:17584-17591(1991).
 FT NON_TER 1 9
 FT NON_TER 9 9
 SQ SEQUENCE 9 AA; 985 MW; 54D92763387DC5B3 CRC64;
 Query Match 28.6%; Score 2; DB 8; Length 9;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 KP 5
 DB 6 KP 7
 RESULT 46
 ID Q62530 PRELIMINARY; PRT; 9 AA.
 AC Q62530;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 DE Osteocalcin (Gamma-carboxyglutamic acid-containing protein) (Bone
 DE GLA-protein) (BGP) (Fragment).
 GN BGLAP1.
 OS Mus spretus (Western wild mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10096;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRET/EL;
 RX MEDLINE=94319082; PubMed=8043949;
 RA Ko M.S., Wang X., Horton J.H., Hagen M.D., Takahashi N., Maezaki Y.,
 RA Nadeau J.H.;
 RT "Genetic mapping of 40 cDNA clones on the mouse genome by PCR.";
 RL Mamm. Genome 5:349-355(1994).
 CC -!- FUNCTION: THIS BONE PROTEIN, CONSTITUTES 1-2% OF THE TOTAL PROTEIN
 CC -!- OF BONE. IT BINDS STRONGLY TO APATITE AND CALCIUM.
 CC -!- TISSUE SPECIFICITY: BONE.
 CC -!- PTM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K
 CC -!- DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE
 CC BINDING OF CALCIUM.
 DR EMBL; U05695; AAB60465.1; -.
 DR MGD; MGI:88156; Bglap1.
 FT NON_TER 1 1
 FT NON_TER 1 1
 SQ SEQUENCE 9 AA; 1135 MW; 5937E0586B504403 CRC64;
 Query Match 28.6%; Score 2; DB 11; Length 9;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 6 YK 7
 DB 1 YK 2

RESULT 47
 Q8CG13 PRELIMINARY; PRT; 9 AA.
 ID Q8CG13;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Glutamate receptor ionotropic N-methyl D-aspartate-like 1A
 DE (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RA Wyther K.S., Mohan Raj B.K., Sciorra L.J., Roginski R.S.;
 RT "The mouse orthologue of the human ionotropic glutamate receptor-like
 RT gene (GRIN1A) maps to mouse chromosome 9.";
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF462417; AAO15648.1; -.
 DR EMBL; AF462416; AAO15648.1; JOINED.
 KW Receptor.
 FT NON_TER 1 1
 FT NON_TER 9 9
 SQ SEQUENCE 9 AA; 1091 MW; 6A91233EB059C33B CRC64;
 Query Match 28.6%; Score 2; DB 11; Length 9;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EK 4
 DB 2 EK 3

RESULT 48
 Q8QRR6 PRELIMINARY; PRT; 9 AA.
 ID Q8QRR6;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Spike glycoprotein (Fragment).
 GN S.
 OS Transmissible gastroenteritis virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Coronaviridae; Coronavirus.
 OX NCBI_TaxID=11149;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HKT2;
 RA Kim S.J., Kwon H.M.;
 RT "Genetic characterization of the spike glycoprotein gene of
 RT transmissible gastroenteritis viruses (TGEVs) isolated in Korea and
 RT the genes 3/3-1 in TGEV passaged on swine testicular cell.";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF481367; AAL69746.1; -.
 FT NON_TER 1 1
 FT NON_TER 1 1
 SQ SEQUENCE 9 AA; 1087 MW; 34E3C1F2C33B1047 CRC64;
 Query Match 28.6%; Score 2; DB 12; Length 9;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EK 4
 DB 4 EK 5

RESULT 49

Q8QRR5 PRELIMINARY; PRT; 9 AA.
 ID Q8QRR5;
 AC Q8QRR5;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Spike glycoprotein (Fragment).
 GN S.
 OS Transmissible gastroenteritis virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Coronaviridae; Coronavirus.
 OX NCBI_TaxID=11149;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=KT2;
 RA Kim S.J., Kwon H.M.;
 RT "Genetic characterization of the spike glycoprotein gene of
 RT transmissible gastroenteritis viruses (TGEV) isolated in Korea and
 RT the genes 3/3.1 in TGEV passaged on swine testicular cell.";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBSJ databases.
 DR EMBL; AF481368; AAL89748.1; -;
 FT NON TER 1
 SQ SEQUENCE 9 AA; 1087 MW; 34E3C1F2C33B1047 CRC64;
 Query Match 28.6%; Score 2; DB 12; Length 9;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 EK 4
 Db 4 EK 5
 RESULT 50
 Q8AUM7 PRELIMINARY; PRT; 9 AA.
 ID Q8AUM7;
 AC Q8AUM7;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Cytochrome P450 aromatase (Fragment).
 GN CYP19A.
 OS Carassius auratus (Goldfish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Carassius.
 OX NCBI_TaxID=7957;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP Tchoudakova A.V., Kishida M., Wood E., Callard G.V.;
 RA "Promoter characteristics of two CYP19 genes differentially expressed
 RT in the brain and ovary of teleost fish.";
 RL J. Steroid Biochem. Mol. Biol. 0:0-0(2001).
 DR EMBL; AF324895; AAN32616.1; -;
 DR EMBL; AF324896; AAN32617.1; -;
 FT NON TER 9
 SQ SEQUENCE 9 AA; 961 MW; C49E76D7272B187D CRC64;
 Query Match 28.6%; Score 2; DB 13; Length 9;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GE 3
 Db 3 GE 4
 RESULT 51
 Q9F9H5 PRELIMINARY; PRT; 10 AA.
 ID Q9F9H5;
 AC Q9F9H5;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE Glutamate racemase (Fragment).
 GN GLR.
 OS Helicobacter pylori (Campylobacter pylori).
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
 OC Helicobacteraceae; Helicobacter.
 OX NCBI_TaxID=210;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=India75A;
 RX MEDLINE=20270152; PubMed=10809702;
 RA Kersulyte D., Mukhopadhyay A.K., Velapattino B., Su W.W., Pan Z.J.,
 Garcia C., Hernandez V., Valdez Y., Mistry R.S., Gilman R.H., Yuan Y.,
 Gao H., Alarcon T., Lopez-Brea M., Balakrishna Nair G., Chowdhury A.,
 Datta S., Shirai M., Nakazawa T., Ally R., Segal I., Wong B.C.,
 Lam S.K., Olfat F.O., Boren T., Engstrand L., Torres O., Schneider R.,
 Thomas J.E., Czinn S., Berg D.E.;
 RA "Differences in genotypes of Helicobacter pylori from different human
 RT populations.";
 RT J. Bacteriol. 182:3210-3218(2000).
 RL J. Bacteriol. 182:3210-3218(2000).
 DR EMBL; AF190663; AAG18486.1; -;
 FT NON TER 1
 SQ SEQUENCE 10 AA; 1273 MW; CAD126337B133DC6 CRC64;
 Query Match 28.6%; Score 2; DB 2; Length 10;
 Best Local Similarity 100.0%; Pred. No. 4.2e+04;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 EK 4
 Db 1 EK 2
 RESULT 52
 Q9JN49 PRELIMINARY; PRT; 10 AA.
 ID Q9JN49;
 AC Q9JN49;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE Plasmid pE194 ermC leader region (Fragment).
 OS Staphylococcus aureus.
 OG Plasmid pE194.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=1280;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP MEDLINE=86041917; PubMed=3903662;
 RA Narayanan C.S., Dubnau D.;
 RA "Evidence for the translational attenuation model: ribosome-binding
 RT studies and structural analysis with an in vitro run-off transcript
 RT of ermC.";
 RL Nucleic Acids Res. 13:7307-7326(1985).
 DR EMBL; X03097; CAA26883.1; -;
 KW Plasmid.
 FT NON TER 10
 SQ SEQUENCE 10 AA; 1228 MW; 6B6C8E33054433B CRC64;
 Query Match 28.6%; Score 2; DB 2; Length 10;
 Best Local Similarity 100.0%; Pred. No. 4.2e+04;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 EK 4
 Db 3 EK 4
 RESULT 53
 Q9XBH3 PRELIMINARY; PRT; 10 AA.
 ID Q9XBH3;
 AC Q9XBH3;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Celf-like protein (Fragment).
 GN Celf.
 OS Bacillus cereus.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1396;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 10987;
 RX MEDLINE=99231848; PubMed=10217496;
 RA Okstad O.A., Hegna I., Lindbaek T., Rishovd A.L., Kolsto A.B.;
 RT "Genome organisation is not conserved between Bacillus cereus and
 RL Bacillus subtilis."
 DR Microbiology 145:621-631(1999).
 FT EMBL; AJ000394; CAB40625.1; --
 NON_TER 1
 SQ SEQUENCE 10 AA; 1264 MW; D3757EC33339C9D6 CRC64;

Query Match 28.6%; Score 2; DB 2; Length 10;
 Best Local Similarity 100.0%; Pred. No. 4.2e+04;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EK 4
 DB 9 EK 10

RESULT 54

Q9ZIB1 PRELIMINARY; PRT; 10 AA.
 AC Q9ZIB1;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
 DE Toxin B (Fragment).
 GN TOXB.
 OS Clostridium difficile.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1496;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=8864;
 RA Dodson A.P., Borriello S.P.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF035716; AAD02038.1; --
 FT NON_TER 1
 SQ SEQUENCE 10 AA; 1120 MW; 913DA042C736DDC1 CRC64;

Query Match 28.6%; Score 2; DB 2; Length 10;
 Best Local Similarity 100.0%; Pred. No. 4.2e+04;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PY 6
 DB 1 PY 2

RESULT 55

Q8KH9 PRELIMINARY; PRT; 10 AA.
 AC Q8KH9;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE BONT/A (Fragment).
 GN BONT/A.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN (1)
 RP SEQUENCE FROM N.A.

RC STRAIN=62A, and NCTC 2916;
 RA Dineen S.S., Bradshaw M., Johnson E.A.;
 RT "Comparison of the neurotoxin gene clusters in Clostridium botulinum
 RT type A strains."
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF461539; AAM75954.1; --
 DR EMBL; AF461541; AAM75962.1; --
 FT NON_TER 1
 SQ SEQUENCE 10 AA; 1143 MW; 8721FA0B1863787A CRC64;

Query Match 28.6%; Score 2; DB 2; Length 10;
 Best Local Similarity 100.0%; Pred. No. 4.2e+04;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GE 3
 DB 6 GE 7

RESULT 56

Q8VN85 PRELIMINARY; PRT; 10 AA.
 ID Q8VN85
 AC Q8VN85;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Urease accessory protein.
 GN UREP.
 OS Helicobacter pylori (Campylobacter pylori).
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 OC Helicobacteraceae; Helicobacter.
 OX NCBI_TaxID=210;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=NQ267;
 RX MEDLINE=21625040; PubMed=11742075;
 RA Falush D., Kraft C., Taylor N.S., Correa P., Fox J.G., Achtman M.,
 RA Suerbaum S.;
 RT "Recombination and mutation during long-term gastric colonization by
 RT Helicobacter pylori: Estimates of clock rates, recombination size and
 RT minimal age."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:15056-15061(2001).
 DR EMBL; AJ418329; CAD11229.1; --
 SQ SEQUENCE 10 AA; 1329 MW; 81E802040332411 CRC64;

Query Match 28.6%; Score 2; DB 2; Length 10;
 Best Local Similarity 100.0%; Pred. No. 4.2e+04;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EK 4
 DB 8 EK 9

RESULT 57

Q47475 PRELIMINARY; PRT; 10 AA.
 ID Q47475
 AC Q47475;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
 DE Pot. pncB protein (Fragment).
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86164315; PubMed=2869947;
 RA Bally M., Foglino M., Bruschi M., Murgier M., Lazdunski A.;
 RT "Nucleotide sequence of the promoter and amino-terminal encoding
 RT region of the Escherichia coli pepN gene."
 RL Eur. J. Biochem. 155:565-569(1986).

DR EMBL; X03709; CAA27335.1; -.

FT NON TER 10 10

SQ SEQUENCE 10 AA; 1094 MW; 04401CBEB05EA5B0 CRC64;

Query Match 28.6%; Score 2; DB 2; Length 10;

Best Local Similarity 100.0%; Pred. No. 4.2e+04;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KP 5

DB 9 KP 10

RESULT 58

O54217

ID O54217 PRELIMINARY; PRT; 10 AA.

AC O54217 (TREMELrel. 06, Created)

DT 01-JUN-1998 (TREMELrel. 06, Created)

DE Neurotoxin type A HN+ 57 kDa subunit (Fragment)

OC Clostridium botulinum.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

OC Clostridium.

OC NCBI_TaxID=1491;

OC RN [1]

OC STaphylococcus epidermidis.

OG Plasmid pCH01.

OC Bacteria; Firmicutes; Bacillales; Staphylococcus.

OC NCBI_TaxID=1282;

OC RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BN 280;

RX MEDLINE=9839475; PubMed=9726851;

RA Heidrich C., Pag U., Josten M., Metzger J., Jack R.W., Bierbaum G.,

RA Jung G., Sahl H.G.;

RT "Isolation, Characterization and sequence of the novel lantibiotic

epicidin 280 and its biosynthetic gene cluster.";

RL Appl. Environ. Microbiol. 64:3140-3146(1998).

DR EMBL; Y14023; CAA74345.1; -.

KW Plasmid.

FT NON TER 10 10

SQ SEQUENCE 10 AA; 1453 MW; 05A0BFD32B4409DB CRC64;

Query Match 28.6%; Score 2; DB 2; Length 10;

Best Local Similarity 100.0%; Pred. No. 4.2e+04;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 YK 7

DB 6 YK 7

RESULT 59

Q9RSN2

ID Q9RSN2 PRELIMINARY; PRT; 10 AA.

AC Q9RSN2 (TREMELrel. 13, Created)

DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)

DE Neurotoxin type A HN+ 57 kDa SUBUNIT-SAMPLE 2 (Fragment).

OC Clostridium botulinum.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

OC Clostridium.

OC NCBI_TaxID=1491;

OC RN [1]

OC SEQUENCE.

RX MEDLINE=92143938; PubMed=1781887;

RA Somers E., DasGupta B.R.;

RT "Clostridium botulinum types A, B, C1, and E produce proteins with or

without hemagglutinating activity: do they share common amino acid

sequences and genes?";

RL J. Protein Chem. 10:415-425(1991).

FT NON TER 1 1

FT NON TER 10 10

SQ SEQUENCE 10 AA; 1225 MW; EC3DE932D366C1BA CRC64;

Query Match 28.6%; Score 2; DB 2; Length 10;

Best Local Similarity 100.0%; Pred. No. 4.2e+04;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PY 6

DB 9 PY 10

RESULT 60

Q9RSN1

ID Q9RSN1 PRELIMINARY; PRT; 10 AA.

AC Q9RSN1 (TREMELrel. 13, Created)

DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)

DE Neurotoxin type A HN+ 57 kDa subunit (Fragment).

OC Clostridium botulinum.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

OC Clostridium.

OC NCBI_TaxID=1491;

OC RN [1]

OC SEQUENCE.

RX MEDLINE=92143938; PubMed=1781887;

RA Somers E., DasGupta B.R.;

RT "Clostridium botulinum types A, B, C1, and E produce proteins with or

without hemagglutinating activity: do they share common amino acid

sequences and genes?";

RL J. Protein Chem. 10:415-425(1991).

FT NON TER 1 1

FT NON TER 10 10

SQ SEQUENCE 10 AA; 1172 MW; E5DAE932D416C1BA CRC64;

Query Match 28.6%; Score 2; DB 2; Length 10;

Best Local Similarity 100.0%; Pred. No. 4.2e+04;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PY 6

DB 9 PY 10

RESULT 61

Q9RSN3

ID Q9RSN3 PRELIMINARY; PRT; 10 AA.

AC Q9RSN3 (TREMELrel. 13, Created)

DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)

DE Neurotoxin type B HN+ 57 kDa subunit (Fragment).

OC Clostridium botulinum.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

OC Clostridium.

OC NCBI_TaxID=1491;

OC RN [1]

OC SEQUENCE.

RX MEDLINE=92143938; PubMed=1781887;

RA Somers E., DasGupta B.R.;

RT "Clostridium botulinum types A, B, C1, and E produce proteins with or

without hemagglutinating activity: do they share common amino acid

sequences and genes?";

RL J. Protein Chem. 10:415-425(1991).

FT NON TER 1 1

FT NON TER 10 10

SQ SEQUENCE 10 AA; 1221 MW; C3DBE932D416C1B2 CRC64;

Query Match 28.6%; Score 2; DB 2; Length 10;

Best Local Similarity 100.0%; Pred. No. 4.2e+04;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PY 6

DB 9 PY 10

```
RESULT 62
Q51812
ID Q51812 PRELIMINARY; PRT; 10 AA.
AC Q51812;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE DNA helicase (Fragment).
GN TRAI.
OS Escherichia coli.
OG Plasmid F.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K-12;
RX MEDLINE=92009201; PubMed=1916281;
RA Cram D.S., Loh S.M., Cheah K.C., Skurray R.A.;
RT "Sequence and conservation of genes at the distal end of the transfer
RT region on plasmids F and R6-S.*";
RL Gene 104:85-90 (1991).
DR EMBL; M38047; AAA98090.1; -.
KW Helicase; Plasmid.
FT NON TER 1
SQ SEQUENCE 10 AA; 1088 MW; 59A3047731A33B13 CRC64;

Query Match 28.6%; Score 2; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.2e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EK 4
DB 4 EK 5

RESULT 63
P83154
ID P83154 PRELIMINARY; PRT; 10 AA.
AC P83154;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE Phycobilisome rod-core linker polypeptide cpcG3 (Fragment).
OS Anabaena sp. (strain L31).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.
OX NCBI_TaxID=29412;
RN [1]
RP SEQUENCE.
RA Apte S.K., Uhlemann E., Schmid R., Altendorf K.;
RL Submitted (OCT-2001) to the SWISS-PROT data bank.
CC -!- FUNCTION: ROD-CORE LINKER PROTEIN REQUIRED FOR ATTACHMENT OF
CC PHYCOCYANIN TO ALLOPHYCOCYANIN IN CORES OF PHYCOBILISOMES.
CC -!- FUNCTION: LINKER POLYPEPTIDES DETERMINE THE STATE OF AGGREGATION
CC AND THE LOCATION OF THE DISK-SHAPED PHYCOBILIPROTEIN UNITS WITHIN
CC THE PHYCOBILISOME AND MODULATE THEIR SPECTROSCOPIC PROPERTIES IN
CC ORDER TO MEDATE A DIRECTED AND OPTIMAL ENERGY TRANSFER.
CC -!- SUBUNIT: THE PHYCOBILISOME IS A HEMIDISCOIDAL STRUCTURE THAT IS
CC COMPOSED OF TWO DISTINCT SUBSTRUCTURES: A CORE COMPLEX (THAT
CC CONTAINS PHYCOBILIPROTEINS) AND A NUMBER OF RODS RADIATING FROM
CC THE CORE.
CC -!- SUBCELLULAR LOCATION: THYLAKOID MEMBRANE.
CC -!- SIMILARITY: TO OTHER PHYCOBILISOME LINKER PROTEINS.
KW Phycobilisome; Photosynthesis; Thylakoid; Membrane.
FT NON TER 10
SQ SEQUENCE 10 AA; 1144 MW; 2F9B662B5B172737 CRC64;

Query Match 28.6%; Score 2; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.2e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EK 4
DB 4 EK 5

RESULT 64
Q9UWV2
ID Q9UWV2 PRELIMINARY; PRT; 10 AA.
AC Q9UWV2;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Bap2(1).
GN BAP2(1).
OS Schizopyllum commune (Bracket fungus).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Agaricales; Schizophyllaceae; Schizophyllum.
OX NCBI_TaxID=5334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12-43;
RA Siebert-Bartholmei C., Kothe E.;
RT "Basidiomycete pheromone recognition studied in yeast.";
RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF102785; AAF03757.1; -.
KW SEQUENCE 10 AA; 1076 MW; D7D803AEB1E861B2 CRC64;

Query Match 28.6%; Score 2; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.2e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
DB 4 TG 5

RESULT 65
Q9UN90
ID Q9UN90 PRELIMINARY; PRT; 10 AA.
AC Q9UN90;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE Canalicular multispecific organic anion transporter (Fragment).
GN CMOAT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tanaka T., Uchiyumi T., Hinoshita E., Inokuchi A., Toh S., Wada M.,
RA Nomoto M., Kohno K., Kuwano M.;
RT "Sequence analysis and functional characterization of the 5'-flanking
RT region of the human canalicular multispecific organic anion
RT transporter/multidrug resistance protein 2 (cMOAT/MRP2) gene.";
RL Submitted (APR-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF144630; RAD47599.1; -.
FT NON TER 10
SQ SEQUENCE 10 AA; 1219 MW; 76F28CB44EB9C33B CRC64;

Query Match 28.6%; Score 2; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.2e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EK 4
DB 3 EK 4

RESULT 66
Q9P229
ID Q9P229 PRELIMINARY; PRT; 10 AA.
```

AC Q9P2Z9;
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
 DE Epidermal growth factor receptor (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=91337091; PubMed=1678600;
 RA Humphrey P.A., Gangarosa L.M., Wong A.J., Archer G.E.,
 RA Lund-Johansen M., Bjerkvig R., Laerum O.D., Friedman H.S.,
 RA Bigner D.D.;
 RT "Deletion-mutant epidermal growth factor receptor in human gliomas:
 RT effects of type II mutation on receptor function.";
 RL Biochem. Biophys. Res. Commun. 178:1413-1420(1991).
 DR EMBL: S51343; AAB19486.2; --
 KW Receptor.
 FT NON_TER 1 1
 FT NON_TER 10 10
 SQ SEQUENCE 10 AA; 1020 MW; 439DEAE8B87B1727 CRC64;
 Query Match 28.6%; Score 2; DB 4; Length 10;
 Best Local Similarity 100.0%; Pred. No. 4.2e+04;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TG 2
 DB 8 TG 9
 RESULT 67
 Q96QT9 PRELIMINARY; PRT; 10 AA.
 AC Q96QT9;
 DT 01-DEC-2001 (TRENBLrel. 19, Created)
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE Protein tyrosine phosphatase 1B (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21326081; PubMed=11316810;
 RA Fukuda T., Tonks N.K.;
 RT "The Reciprocal Role of Egr-1 and Sp Family Proteins in Regulation of
 RT the PTPB Promoter in Response to the p210 Bcr-Abl Oncoprotein-
 RT tyrosine Kinase.";
 RL J. Biol. Chem. 276:25512-25519(2001).
 DR EMBL: AY029236; AAK31734.1; --
 FT NON_TER 10 10
 FT NON_TER 10 10
 SQ SEQUENCE 10 AA; 1314 MW; AC08E219CB133B16 CRC64;
 Query Match 28.6%; Score 2; DB 4; Length 10;
 Best Local Similarity 100.0%; Pred. No. 4.2e+04;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 EK 4
 DB 4 EK 5
 RESULT 68
 Q8WXB5 PRELIMINARY; PRT; 10 AA.
 AC Q8WXB5;
 DT 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)

DE EYAZB (Fragment).
 GN EYAZ.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Fee B.E., Doyle C.A., Cleveland J.L.;
 RT "A novel eyes absent protein is expressed in the human eye.";
 RL Gene 0:0-0(2002).
 DR EMBL: AF455148; AAL57875.1; --
 FT NON_TER 10 10
 FT NON_TER 10 10
 SQ SEQUENCE 10 AA; 1298 MW; 18021202C69B132B CRC64;
 Query Match 28.6%; Score 2; DB 4; Length 10;
 Best Local Similarity 100.0%; Pred. No. 4.2e+04;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 6 YK 7
 DB 3 YK 4
 RESULT 69
 Q81ZA2 PRELIMINARY; PRT; 10 AA.
 AC Q81ZA2;
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Cardiac troponin T (Fragment).
 GN TNNT2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RA Bahl A., Kubo T., Steffensen U., Steffensen M., McKenna W.J.,
 RA Mogensen J.;
 RT "Homo sapiens Troponin T (TNNT2) introns 6 and 7 and exon 7.";
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY160215; AAN71650.1; --
 FT NON_TER 1 1
 FT NON_TER 10 10
 SQ SEQUENCE 10 AA; 1099 MW; AB24BC6325BB1B:6 CRC64;
 Query Match 28.6%; Score 2; DB 4; Length 10;
 Best Local Similarity 100.0%; Pred. No. 4.2e+04;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 KP 5
 DB 7 KP 8
 RESULT 70
 Q9TR47 PRELIMINARY; PRT; 10 AA.
 AC Q9TR47;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE Amphoterin homolog (Fragment).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE.

```

RX MEDLINE=96029671; PubMed=7592757;
RA Hori O., Brett J., Slattery T., Cao R., Zhang J., Chen J.X.,
RA Nagashima M., Lundh E.R., Vijay S., Nitecki D.;
RT "The receptor for advanced glycation end products (RAGE) is a cellular
RT binding site for amphoterin. Mediation of neurite outgrowth and co-
RT expression of rage and amphoterin in the developing nervous system.";
RL J. Biol. Chem. 270:25752-25761(1995)
SQ SEQUENCE 10 AA; 1163 MW; 28E5034453769B18 CRC64;

Query Match      28.6%; Score 2; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.2e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 GE 3
DB      3 GE 4

RESULT 71
ID Q9TR48      PRELIMINARY;      PRT;      10 AA.
AC Q9TR48;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Amphoterin (Bovine).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=96029671; PubMed=7592757;
RA Hori O., Brett J., Slattery T., Cao R., Zhang J., Chen J.X.,
RA Nagashima M., Lundh E.R., Vijay S., Nitecki D.;
RT "The receptor for advanced glycation end products (RAGE) is a cellular
RT binding site for amphoterin. Mediation of neurite outgrowth and co-
RT expression of rage and amphoterin in the developing nervous system.";
RL J. Biol. Chem. 270:25752-25761(1995)
SQ SEQUENCE 10 AA; 1050 MW; 23B89A286761EB18 CRC64;

Query Match      28.6%; Score 2; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.2e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 GE 3
DB      3 GE 4

RESULT 72
ID Q8H271      PRELIMINARY;      PRT;      10 AA.
AC Q8H271;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Alpha 1B glycoprotein DVOP32-1 (Fragment).
OS Didelphis marsupialis virginiana (North American opossum).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Didelphimorphia; Didelphidae; Didelphis.
OX NCBI_TaxID=9267;
RN [1]
RP SEQUENCE FROM N.A.
RX Martinez M.E., Pierce J.R.;
RA "Molecular Characterization and Expression of an Antihemorrhagic Gene
RA Family from Didelphis virginiana.";
RT Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY131004; AAN06915.1; -.
FT NON_TER
SQ SEQUENCE 10 AA; 1072 MW; 69FE4CB1B2D2C72B CRC64;

```

```

Query Match      28.6%; Score 2; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.2e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TG 2
DB      8 TG 9

RESULT 73
ID Q8MAZ9      PRELIMINARY;      PRT;      10 AA.
AC Q8MAZ9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE PsbJ (Fragment).
GN PSBJ.
OS Dicranostyles ampla.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; Lamiales; Solanales; Convolvulaceae; Dicranostyles.
OX NCBI_TaxID=197378;
RN [1]
RP SEQUENCE FROM N.A.
RA Stefanovic S., Krueger L., Olmstead R.G.;
RT "Monophyly of the Convolvulaceae and circumscription of their major
RT lineages based on DNA sequences of multiple chloroplast loci.";
RL Am. J. Bot. 0:0-0(2002).
DR EMBL; AY100933; AAM55853.1; -.
KW Chloroplast.
FT NON_TER
SQ SEQUENCE 10 AA; 1074 MW; 836D59440861B1AA CRC64;

Query Match      28.6%; Score 2; DB 8; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.2e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TG 2
DB      5 TG 6

RESULT 74
ID Q8MBB7      PRELIMINARY;      PRT;      10 AA.
AC Q8MBB7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE PsbJ (Fragment).
GN PSBJ.
OS Merremia aegyptia.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; Lamiales; Solanales; Convolvulaceae; Merremia.
OX NCBI_TaxID=197413;
RN [1]
RP SEQUENCE FROM N.A.
RA Stefanovic S., Krueger L., Olmstead R.G.;
RT "Monophyly of the Convolvulaceae and circumscription of their major
RT lineages based on DNA sequences of multiple chloroplast loci.";
RL Am. J. Bot. 0:0-0(2002).
DR EMBL; AY100875; AAM55624.1; -.
KW Chloroplast.
FT NON_TER
SQ SEQUENCE 10 AA; 1074 MW; 836D59440861B1AA CRC64;

Query Match      28.6%; Score 2; DB 8; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.2e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 TG 2
DB      5 TG 6

RESULT 75
P82937
ID P82937 PRELIMINARY; PRT; 10 AA.
AC
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE Unknown endospore protein B (Fragment).
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Triticeae; Hordeum.
OX NCB1_TaxID=4513;
RN [1]
RP SEQUENCE.
RC STRAIN=cv. BOM1; TISSUE=Starchy endospore;
RX MEDLINE=2108911; PubMed=11271488;
RA Kristoffersen H.E., Flengsrud R.;
RT "Separation and characterization of basic barley seed proteins.";
RL Electrophoresis 21:3693-3700(2000).
CC -:- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 8.5-9.0, ITS MW IS: 11.9 KDA.
FT NON TER 10
SQ SEQUENCE 10 AA; 1297 MW; 8248A50B11PB5EBA CRC64;

Query Match 28.6%; Score 2; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.2e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 EK 4
DB      1 EK 2

RESULT 76
Q9QVK8
ID Q9QVK8 PRELIMINARY; PRT; 10 AA.
AC
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE MEPRIN-METALLOENDOPETIDASE (Fragment).
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCB1_TaxID=10095;
RN [1]
RP SEQUENCE.
RX MEDLINE=91363409; PubMed=1888759;
RA Flannery A.V., Macadam G.C., Beynon R.J.;
RT "Immunological characterisation of different meprin species in mice.";
RL Biochim. Biophys. Acta 1079:119-122(1991).
FT NON TER 1 1
FT NON TER 10 10
SQ SEQUENCE 10 AA; 1163 MW; DD6436144731B2C9 CRC64;

Query Match 28.6%; Score 2; DB 11; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.2e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 EK 4
DB      9 EK 10

RESULT 77
Q9R0Z0
ID Q9R0Z0 PRELIMINARY; PRT; 10 AA.
AC
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DE CW17 protein (Fragment).
GN CW17.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCB1_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Liver;
RX MEDLINE=99287587; PubMed=10360842;
RA Wrehlke C., Wiedemeyer W.R., Schmitt-Wrede H.P., Mincheva A.,
RA Lichter P., Wunderlich F.;
RT "Genomic organization of mouse gene zfp162 (mzfml).";
RL DNA Cell Biol. 18:419-428(1999).
DR EMBL; Y14702; CAB45189.1; -.
FT NON TER 10 10
SQ SEQUENCE 10 AA; 946 MW; 836D48ADD44D861 CRC64;

Query Match 28.6%; Score 2; DB 11; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.2e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TG 2
DB      3 TG 4

RESULT 78
Q9Q0W9
ID Q9Q0W9 PRELIMINARY; PRT; 10 AA.
AC
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE Large T antigen (Fragment).
OS Polyomavirus JC.
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCB1_TaxID=10632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JME43;
RX MEDLINE=20087544; PubMed=10618230;
RA Bofill-Mas S., Fina S., Girones R.;
RT "Documenting the epidemiologic patterns of polyomaviruses in human
RT populations by studying their presence in urban sewage.";
RL Appl. Environ. Microbiol. 66:238-245(2000).
DR EMBL; AF119350; AAF24106.1; -.
FT NON TER 1 1
SQ SEQUENCE 10 AA; 1167 MW; 4CD6A97771A32763 CRC64;

Query Match 28.6%; Score 2; DB 12; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.2e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 KP 5
DB      3 KP 4

RESULT 79
Q8JV70
ID Q8JV70 PRELIMINARY; PRT; 10 AA.
AC
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Large T antigen (Fragment).
OS Polyomavirus JC.

```


OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USA 5;
RX MEDLINE=21465052; PubMed=11581397;
RA Bofill-Mas S., Formiga-Cruz M., Clemente-Casares P., Calafell F.,
Girones R.;
RT "Potential transmission of human polyomaviruses through the
gastrointestinal tract after exposure to virions or viral DNA.";
RL J. Virol. 75:10290-10299(2001).
DR EMBL; AF304387; AAM97804.1; -.
FT NON_TER 1
SQ SEQUENCE 10 AA; 1167 MW; 4CD6A97771A32763 CRC64;

Query Match 28.6%; Score 2; DB 12; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.2e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KP 5
Db 3 KP 4

RESULT 80

O9Q0W1 ID Q9Q0W1 PRELIMINARY; PRT; 10 AA.
AC Q9Q0W1
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE Large T antigen (Fragment).
OS Polyomavirus JC.
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CSFB;
RX MEDLINE=20087544; PubMed=10618230;
RA Bofill-Mas S., Pina S., Girones R.;
RT "Documenting the epidemiologic patterns of polyomaviruses in human
populations by studying their presence in urban sewage.";
RL Appl. Environ. Microbiol. 66:238-245(2000).
DR EMBL; AF119354; AAF24114.1; -.
FT NON_TER 1
SQ SEQUENCE 10 AA; 1167 MW; 4CD6A97771A32763 CRC64;

Query Match 28.6%; Score 2; DB 12; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.2e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KP 5
Db 3 KP 4

RESULT 81

O8JUV68 ID O8JUV68 PRELIMINARY; PRT; 10 AA.
AC O8JUV68
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Large T antigen (Fragment).
OS Polyomavirus JC.
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USA 6;
RX MEDLINE=21465052; PubMed=11581397;
RA Bofill-Mas S., Formiga-Cruz M., Clemente-Casares P., Calafell F.,
Girones R.;

RT "Potential transmission of human polyomaviruses through the
gastrointestinal tract after exposure to virions or viral DNA.";
RL J. Virol. 75:10290-10299(2001).
DR EMBL; AF304388; AAM97806.1; -.
FT NON_TER 1
SQ SEQUENCE 10 AA; 1167 MW; 4CD6A97771A32763 CRC64;

Query Match 28.6%; Score 2; DB 12; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.2e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KP 5
Db 3 KP 4

RESULT 82

O9Q0V9 ID O9Q0V9 PRELIMINARY; PRT; 10 AA.
AC O9Q0V9
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE Large T antigen (Fragment).
OS Polyomavirus JC.
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CSFB;
RX MEDLINE=20087544; PubMed=10618230;
RA Bofill-Mas S., Pina S., Girones R.;
RT "Documenting the epidemiologic patterns of polyomaviruses in human
populations by studying their presence in urban sewage.";
RL Appl. Environ. Microbiol. 66:238-245(2000).
DR EMBL; AF119355; AAF24116.1; -.
FT NON_TER 1
SQ SEQUENCE 10 AA; 1167 MW; 4CD6A97771A32763 CRC64;

Query Match 28.6%; Score 2; DB 12; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.2e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KP 5
Db 3 KP 4

RESULT 83

O90349 ID O90349 PRELIMINARY; PRT; 10 AA.
AC O90349
DT 01-NOV-1998 (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis G virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OX GBV-C/HGV group.
OX NCBI_TaxID=45255;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SG3362;
RX MEDLINE=99266893; PubMed=10335862;
RA Wong S.B.J., Chan S.H., Ren E.C.;
RT "Diversity of GB virus C/hepatitis G virus isolates in Singapore:
predominance of group 2a and the Asian group 3 variant.";
RL J. Med. Virol. 58:145-153(1999).
DR EMBL; AF078060; AAC32367.1; -.
FT NON_TER 10
SQ SEQUENCE 10 AA; 1147 MW; 623E3E272DCB4774 CRC64;

Query Match 28.6%; Score 2; DB 12; Length 10;

Best Local Similarity 100.0%; Pred. No. 4.2e+04; Mismatches 0; Indels 0; Gaps 0;

QY 5 PY 6
DB 4 PY 5

RESULT 84

Q8JV84 ID Q8JV84 PRELIMINARY; PRT; 10 AA.
AC Q8JV84;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DE Large T antigen (Fragment).
OS Polyomavirus JC.
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=G1;
RX MEDLINE=21465052; PubMed=11581397;
RA Bofill-Mas S., Formiga-Cruz M., Clemente-Casares P., Calafell F.,
RA Girones R.;
RT "Potential transmission of human polyomaviruses through the
RT gastrointestinal tract after exposure to virions or viral DNA.";
RL J. Virol. 75:10290-10299(2001).
DR EMBL; AF303943; AAM97790.1; -.
FT NON TER 1
SQ SEQUENCE 10 AA; 1180 MW; 4CD6A97771A32763 CRC64;

Query Match 28.6%; Score 2; DB 12; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.2e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KP 5
DB 3 KP 4

RESULT 85

Q9Q0W7 ID Q9Q0W7 PRELIMINARY; PRT; 10 AA.
AC Q9Q0W7;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE Large T antigen (Fragment).
OS Polyomavirus JC.
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NANCY2;
RX MEDLINE=20087544; PubMed=10618230;
RA Bofill-Mas S., Pina S., Girones R.;
RT "Documenting the epidemiologic patterns of polyomaviruses in human
RT populations by studying their presence in urban sewage.";
RL Appl. Environ. Microbiol. 66:238-245(2000).
DR EMBL; AF119351; AAF24108.1; -.
FT NON TER 1
SQ SEQUENCE 10 AA; 1167 MW; 4CD6A97771A32763 CRC64;

Query Match 28.6%; Score 2; DB 12; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.2e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KP 5
DB 3 KP 4

RESULT 86

Q8JV66 ID Q8JV66 PRELIMINARY; PRT; 10 AA.
AC Q8JV66;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Large T antigen (Fragment).
OS Polyomavirus JC.
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CSF K;
RX MEDLINE=21465052; PubMed=11581397;
RA Bofill-Mas S., Formiga-Cruz M., Clemente-Casares P., Calafell F.,
RA Girones R.;
RT "Potential transmission of human polyomaviruses through the
RT gastrointestinal tract after exposure to virions or viral DNA.";
RL J. Virol. 75:10290-10299(2001).
DR EMBL; AF304389; AAM97808.1; -.
FT NON TER 1
SQ SEQUENCE 10 AA; 1167 MW; 4CD6A97771A32763 CRC64;

Query Match 28.6%; Score 2; DB 12; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.2e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KP 5
DB 3 KP 4

RESULT 87

Q9Q0V7 ID Q9Q0V7 PRELIMINARY; PRT; 10 AA.
AC Q9Q0V7;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE Large T antigen (Fragment).
OS Polyomavirus JC.
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CSFJ;
RX MEDLINE=20087544; PubMed=10618230;
RA Bofill-Mas S., Pina S., Girones R.;
RT "Documenting the epidemiologic patterns of polyomaviruses in human
RT populations by studying their presence in urban sewage.";
RL Appl. Environ. Microbiol. 66:238-245(2000).
DR EMBL; AF119356; AAF24118.1; -.
FT NON TER 1
SQ SEQUENCE 10 AA; 1167 MW; 4CD6A97771A32763 CRC64;

Query Match 28.6%; Score 2; DB 12; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.2e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KP 5
DB 3 KP 4

RESULT 88

Q8JV82 ID Q8JV82 PRELIMINARY; PRT; 10 AA.
AC Q8JV82;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Large T antigen (Fragment).

OS Polyomavirus JC.
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E1;
RX MEDLINE=21465052; PubMed=11581397;
RA Bofill-Mas S., Formiga-Cruz M., Clemente-Casares P., Calafell F.,
RA Girones R.;
RT "Potential transmission of human polyomaviruses through the
RT gastrointestinal tract after exposure to virions or viral DNA.";
RL J. Virol. 75:10290-10299(2001).
DR EMBL; AF303944; AAM97792.1; -.
FT NON_TER 1
SQ SEQUENCE 10 AA; 1167 MW; 4CD6A97771A32763 CRC64;

Query Match 28.6%; Score 2; DB 12; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.2e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KP 5
Db 3 KP 4

RESULT 89
OSJUV76 PRELIMINARY; PRT; 10 AA.
ID Q8JV76
AC Q8JV76; 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Large T antigen (Fragment).
OS Polyomavirus JC.
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USA2;
RX MEDLINE=21465052; PubMed=11581397;
RA Bofill-Mas S., Formiga-Cruz M., Clemente-Casares P., Calafell F.,
RA Girones R.;
RT "Potential transmission of human polyomaviruses through the
RT gastrointestinal tract after exposure to virions or viral DNA.";
RL J. Virol. 75:10290-10299(2001).
DR EMBL; AF303947; AAM97798.1; -.
FT NON_TER 1
SQ SEQUENCE 10 AA; 1167 MW; 4CD6A97771A32763 CRC64;

Query Match 28.6%; Score 2; DB 12; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.2e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KP 5
Db 3 KP 4

RESULT 90
OSJUV74 PRELIMINARY; PRT; 10 AA.
ID Q8JV74
AC Q8JV74; 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Large T antigen (Fragment).
OS Polyomavirus JC.
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USA3;
RX MEDLINE=21465052; PubMed=11581397;

RA Bofill-Mas S., Formiga-Cruz M., Clemente-Casares P., Calafell F.,
RA Girones R.;
RT "Potential transmission of human polyomaviruses through the
RT gastrointestinal tract after exposure to virions or viral DNA.";
RL J. Virol. 75:10290-10299(2001).
DR EMBL; AF303948; AAM97800.1; -.
FT NON_TER 1
SQ SEQUENCE 10 AA; 1167 MW; 4CD6A97771A32763 CRC64;

Query Match 28.6%; Score 2; DB 12; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.2e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KP 5
Db 3 KP 4

RESULT 91
OSJUV77 PRELIMINARY; PRT; 10 AA.
ID Q9Q0X7
AC Q9Q0X7; 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE Large T antigen (Fragment).
OS Polyomavirus JC.
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCN2;
RX MEDLINE=20087544; PubMed=10618230;
RA Bofill-Mas S., Pina S., Girones R.;
RT "Documenting the epidemiologic patterns of polyomaviruses in human
RT populations by studying their presence in urban sewage.";
RL Appl. Environ. Microbiol. 66:238-245(2000).
DR EMBL; AF119346; AAF24098.1; -.
FT NON_TER 1
SQ SEQUENCE 10 AA; 1180 MW; 4CD6A4F771A32763 CRC64;

Query Match 28.6%; Score 2; DB 12; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.2e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KP 5
Db 3 KP 4

RESULT 92
OSJUV78 PRELIMINARY; PRT; 10 AA.
ID Q8JV78
AC Q8JV78; 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Large T antigen (Fragment).
OS Polyomavirus JC.
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USA1;
RX MEDLINE=21465052; PubMed=11581397;
RA Bofill-Mas S., Formiga-Cruz M., Clemente-Casares P., Calafell F.,
RA Girones R.;
RT "Potential transmission of human polyomaviruses through the
RT gastrointestinal tract after exposure to virions or viral DNA.";
RL J. Virol. 75:10290-10299(2001).
DR EMBL; AF303946; AAM97796.1; -.
FT NON_TER 1
SQ SEQUENCE 10 AA; 1183 MW; 4CD6A96731A32763 CRC64;

Query Match 28.6%; Score 2; DB 12; Length 10;
 Best Local Similarity 100.0%; Pred. No. 4.2e+04;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KP 5
 Db 3 KP 4

RESULT 93
 Q9Q0W5 ID Q9Q0W5 PRELIMINARY; PRT; 10 AA.
 AC Q9Q0W5 01-MAY-2000 (TREMELrel. 13, Created)
 DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMELrel. 13, Last annotation update)
 DE Large T antigen (Fragment).
 OS Polyomavirus JC.
 OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
 OX NCBI_TaxID=10632;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BCN8;
 RX Bofill-Mas S., Pina S., Girones R.;
 RA "Documenting the epidemiologic patterns of polyomaviruses in human
 RT populations by studying their presence in urban sewage.";
 RL Appl. Environ. Microbiol. 66:238-245(2000).
 DR EMBL; AF119352; AAF24110.1; -.
 FT NON TER 1
 SQ SEQUENCE 10 AA; 1167 MW; 4CD6A97771A32763 CRC64;

Query Match 28.6%; Score 2; DB 12; Length 10;
 Best Local Similarity 100.0%; Pred. No. 4.2e+04;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KP 5
 Db 3 KP 4

RESULT 94
 Q9Q0X3 ID Q9Q0X3 PRELIMINARY; PRT; 10 AA.
 AC Q9Q0X3 01-MAY-2000 (TREMELrel. 13, Created)
 DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMELrel. 13, Last annotation update)
 DE Large T antigen (Fragment).
 OS Polyomavirus JC.
 OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
 OX NCBI_TaxID=10632;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BCN8;
 RX Bofill-Mas S., Pina S., Girones R.;
 RA "Documenting the epidemiologic patterns of polyomaviruses in human
 RT populations by studying their presence in urban sewage.";
 RL Appl. Environ. Microbiol. 66:238-245(2000).
 DR EMBL; AF119348; AAF24102.1; -.
 FT NON TER 1
 SQ SEQUENCE 10 AA; 1167 MW; 4CD6A97771A32763 CRC64;

Query Match 28.6%; Score 2; DB 12; Length 10;
 Best Local Similarity 100.0%; Pred. No. 4.2e+04;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KP 5
 Db 3 KP 4

RESULT 95
 Q9Q0X5 ID Q9Q0X5 PRELIMINARY; PRT; 10 AA.
 AC Q9Q0X5 01-MAY-2000 (TREMELrel. 13, Created)
 DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMELrel. 13, Last annotation update)
 DE Large T antigen (Fragment).
 OS Polyomavirus JC.
 OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
 OX NCBI_TaxID=10632;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BCN16;
 RX MEDLINE=20087544; PubMed=10618230;
 RA Bofill-Mas S., Pina S., Girones R.;
 RT "Documenting the epidemiologic patterns of polyomaviruses in human
 RT populations by studying their presence in urban sewage.";
 RL Appl. Environ. Microbiol. 66:238-245(2000).
 DR EMBL; AF119347; AAF24100.1; -.
 FT NON TER 1
 SQ SEQUENCE 10 AA; 1167 MW; 4CD6A97771A32763 CRC64;

Query Match 28.6%; Score 2; DB 12; Length 10;
 Best Local Similarity 100.0%; Pred. No. 4.2e+04;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KP 5
 Db 3 KP 4

RESULT 96
 Q9Q0W3 ID Q9Q0W3 PRELIMINARY; PRT; 10 AA.
 AC Q9Q0W3 01-MAY-2000 (TREMELrel. 13, Created)
 DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMELrel. 13, Last annotation update)
 DE Large T antigen (Fragment).
 OS Polyomavirus JC.
 OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
 OX NCBI_TaxID=10632;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PRETORIA1;
 RX MEDLINE=20087544; PubMed=10618230;
 RA Bofill-Mas S., Pina S., Girones R.;
 RT "Documenting the epidemiologic patterns of polyomaviruses in human
 RT populations by studying their presence in urban sewage.";
 RL Appl. Environ. Microbiol. 66:238-245(2000).
 DR EMBL; AF119353; AAF24112.1; -.
 FT NON TER 1
 SQ SEQUENCE 10 AA; 1167 MW; 4CD6A97771A32763 CRC64;

Query Match 28.6%; Score 2; DB 12; Length 10;
 Best Local Similarity 100.0%; Pred. No. 4.2e+04;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KP 5
 Db 3 KP 4

RESULT 97
 Q8JV80 ID Q8JV80 PRELIMINARY; PRT; 10 AA.
 AC Q8JV80 01-OCT-2002 (TREMELrel. 22, Created)
 DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
 DT 01-OCT-2002 (TREMELrel. 22, Last annotation update)
 DE Large T antigen (Fragment).

OS Polyomavirus JC.
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E2;
RX MEDLINE=21465052; PubMed=11581397;
RA Bofill-Mas S., Formiga-Cruz M., Clemente-Casares P., Calafell F.,
RA Girones R.;
RT "Potential transmission of human polyomaviruses through the
RT gastrointestinal tract after exposure to virions of viral DNA.";
RL J. Virol. 75:10290-10299(2001).
DR EMBL; AF303945; AAM97794.1; -.
FT NON TER 1
SQ SEQUENCE 10 AA; 1167 MW; 4CD6A97771A32763 CRC64;

Query Match 28.6%; Score 2; DB 12; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.2e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KP 5
Db 3 KP 4

RESULT 98
Q9Q0X1 PRELIMINARY; PRT; 10 AA.
ID Q9Q0X1
AC Q9Q0X1
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE Large T antigen (Fragment).
OS Polyomavirus JC.
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCN15;
RX MEDLINE=20087544; PubMed=10618230;
RA Bofill-Mas S., Pina S., Girones R.;
RT "Documenting the epidemiologic patterns of polyomaviruses in human
RT populations by studying their presence in urban sewage.";
RL Appl. Environ. Microbiol. 66:238-245(2000).
DR EMBL; AF119349; AAF24104.1; -.
FT NON TER 1
SQ SEQUENCE 10 AA; 1167 MW; 4CD6A97771A32763 CRC64;

Query Match 28.6%; Score 2; DB 12; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.2e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KP 5
Db 3 KP 4

RESULT 99
Q9Q0X9 PRELIMINARY; PRT; 10 AA.
ID Q9Q0X9
AC Q9Q0X9
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE Large T antigen (Fragment).
OS Polyomavirus JC.
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCNU;
RX MEDLINE=20087544; PubMed=10618230;
RA Bofill-Mas S., Pina S., Girones R.;

RT "Documenting the epidemiologic patterns of polyomaviruses in human
RT populations by studying their presence in urban sewage.";
RL Appl. Environ. Microbiol. 66:238-245(2000).
DR EMBL; AF119345; AAF24096.1; -.
FT NON TER 1

SQ SEQUENCE 10 AA; 1167 MW; 4CD6A97771A32763 CRC64;
Query Match 28.6%; Score 2; DB 12; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.2e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KP 5
Db 3 KP 4

RESULT 100
Q8JV72 PRELIMINARY; PRT; 10 AA.
ID Q8JV72
AC Q8JV72
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Large T antigen (Fragment).
OS Polyomavirus JC.
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USA 4;
RX MEDLINE=21465052; PubMed=11581397;
RA Bofill-Mas S., Formiga-Cruz M., Clemente-Casares P., Calafell F.,
RA Girones R.;
RT "Potential transmission of human polyomaviruses through the
RT gastrointestinal tract after exposure to virions of viral DNA.";
RL J. Virol. 75:10290-10299(2001).
DR EMBL; AF304386; AAM97802.1; -.
FT NON TER 1
SQ SEQUENCE 10 AA; 1167 MW; 4CD6A97771A32763 CRC64;

Query Match 28.6%; Score 2; DB 12; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.2e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KP 5
Db 3 KP 4

Search completed: October 21, 2003, 18:27:04
Job time : 67.6111 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:12:30 : Search time 12.0556 Seconds
(without alignments)
27.306 Million cell updates/sec

Title: US-10-057-890A-8

Perfect score: 7

Sequence: 1 *GEKPYK 7

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 372

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Listing first 150 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2	28.6	4	1 TUPT_HUMAN	P01858 homo sapien
2	2	28.6	6	1 OWM_LEPDE	P42985 leptinotars
3	2	28.6	7	1 ALL3_CARMA	P81806 carcinus ma
4	2	28.6	7	1 ALL4_CARMA	P81807 carcinus ma
5	2	28.6	7	1 ALL5_CARMA	P81808 carcinus ma
6	2	28.6	7	1 FAR4_PANRE	P41875 panagrellus
7	2	28.6	7	1 GRFP_MOUSE	P99025 mus musculu
8	2	28.6	7	1 HV7_FIG	P01153 sus scrofa
9	2	28.6	7	1 TPFY_PACDA	P83455 pachymedusa
10	2	28.6	7	1 UF04_MOUSE	P38642 mus musculu
11	2	28.6	8	1 AKHG_GRYBI	P14086 gryllus bim
12	2	28.6	8	1 AL15_CARMA	P81818 carcinus ma
13	2	28.6	8	1 AL16_CARMA	P81819 carcinus ma
14	2	28.6	8	1 ALL3_CYPVO	P82154 cydia pomon
15	2	28.6	8	1 ALL4_CALPO	P41840 calliphora
16	2	28.6	8	1 ALL4_CYPVO	P82155 cydia pomon
17	2	28.6	8	1 ALL5_CALVO	P41841 calliphora
18	2	28.6	8	1 ALL7_CARMA	P81809 carcinus ma
19	2	28.6	8	1 ALL8_CARMA	P81811 carcinus ma
20	2	28.6	8	1 ALL9_CARMA	P81812 carcinus ma
21	2	28.6	8	1 B44K_PORGI	P81886 porphyronon
22	2	28.6	8	1 ORMY_ORCLI	P82455 orconectes
23	2	28.6	8	1 RS1_ERWCH	P37985 erwina chr
24	2	28.6	9	1 AL10_CARMA	P81813 carcinus ma
25	2	28.6	9	1 AL11_CARMA	P81814 carcinus ma
26	2	28.6	9	1 BUK_CLOPA	P81337 clostridium
27	2	28.6	9	1 CCAP_CARMA	P38556 carcinus ma
28	2	28.6	9	1 COW_CONVE	P83047 conus ventr
29	2	28.6	9	1 COXE_THUOE	P80975 thunnus obe
30	2	28.6	9	1 DISIP_RABIT	P01158 oryctolagus
31	2	28.6	9	1 FAR4_CALVO	P41859 calliphora
32	2	28.6	9	1 FAR5_ASCSU	P43170 ascaris suu
33	2	28.6	9	1 FLA2_TREHY	P80159 treponema h

34	2	28.6	9	1 LITR_PHYRO	P08946 phyllomedus
35	2	28.6	9	1 NEUX_HUMAN	P04277 homo sapien
36	2	28.6	9	1 OXYT_EISFO	P42998 eisenia foe
37	2	28.6	9	1 ULAE_HUMAN	P31931 homo sapien
38	2	28.6	9	1 UPAT_HUMAN	P30093 homo sapien
39	2	28.6	10	1 CAL2_LITCI	P82086 litoria cit
40	2	28.6	10	1 CASR_LITXA	P56264 litoria xan
41	2	28.6	10	1 COXQ_RABIT	P80336 oryctolagus
42	2	28.6	10	1 COXQ_SHEEP	P80337 ovis aries
43	2	28.6	10	1 GONI_CHEPR	P80677 chelyosoma
44	2	28.6	10	1 GONI_PETMA	P04378 petromyzon
45	2	28.6	10	1 GONI_PETMA	P30948 petromyzon
46	2	28.6	10	1 PNEU_HUMAN	P22103 homo sapien
47	2	28.6	10	1 PNEU_RAT	P21996 rattus norv
48	2	28.6	10	1 PPCK_FASHE	P80525 fasciola he
49	2	28.6	10	1 QZOG_COMTE	P80466 comamonas t
50	2	28.6	10	1 RT02_BOVIN	P82923 bos taurus
51	2	28.6	10	1 SPI_HALRO	Q10997 halocynthia
52	2	28.6	10	1 TKNB_RANCA	P22689 rana cateab
53	2	28.6	10	1 TKS1_AEDAE	P42634 aedes aegyp
54	2	28.6	10	1 TKS2_AEDAE	P42635 aedes aegyp
55	2	28.6	10	1 URA1_HUMAN	P32118 homo sapien
56	2	28.6	10	1 URA7_HUMAN	P34990 homo sapien
57	2	28.6	10	1 URE3_MORMO	P17339 morganelia
58	2	28.6	10	1 VEG6_BACSU	P80699 bacillus su
59	1	14.3	3	1 GRWV_HUMAN	P01157 homo sapien
60	1	14.3	3	1 LUXE_VIBFI	P24272 vibrio fisc
61	1	14.3	3	1 THYL_PIG	P01151 sus scrofa
62	1	14.3	4	1 ACH1_ACHFU	P35904 achatina fu
63	1	14.3	4	1 DCML_PSECH	P19916 pseudomonas
64	1	14.3	4	1 DCMS_PSECH	P19918 pseudomonas
65	1	14.3	4	1 EOSI_HUMAN	P02731 homo sapien
66	1	14.3	4	1 FAR3_HIRME	P42562 hirudo medi
67	1	14.3	4	1 FAR4_HIRME	P42563 hirudo medi
68	1	14.3	4	1 FFKA_ANTEL	P58705 anthopleura
69	1	14.3	4	1 FYRI_ANTEL	P58706 anthopleura
70	1	14.3	4	1 OCP1_OCTMI	P58648 octopus min
71	1	14.3	4	1 OCP3_OCTMI	P58649 octopus min
72	1	14.3	4	1 RM01_YEAST	P36515 saccharomyc
73	1	14.3	5	1 ALI4_CARMA	P81817 carcinus ma
74	1	14.3	5	1 B10A_CITFR	P13071 citrobacter
75	1	14.3	5	1 BPP7_BOTIN	P30425 bothrops in
76	1	14.3	5	1 EI03_LITRU	P82099 litoria rub
77	1	14.3	5	1 EI04_LITRU	P82100 litoria rub
78	1	14.3	5	1 FARP_ARTTR	P41853 artiopesthi
79	1	14.3	5	1 FARP2_PARMA	P81864 parachitrus
80	1	14.3	5	1 PRCT_FERAM	P01373 periplaneta
81	1	14.3	5	1 PSK_DAUCA	P58261 daucus caro
82	1	14.3	5	1 RE21_LITRU	P82071 litoria rub
83	1	14.3	5	1 RE31_LITRU	P82072 litoria rub
84	1	14.3	5	1 RE32_LITRU	P82073 litoria rub
85	1	14.3	5	1 SUGA_ACHDO	P19991 acheta dome
86	1	14.3	5	1 TPIS_CANPA	P54714 canis famil
87	1	14.3	5	1 TRM3_ECOLI	P13973 escherichia
88	1	14.3	5	1 UC22_MAIZE	P38628 zea mays (m
89	1	14.3	5	1 UF01_MOUSE	P38639 mus musculu
90	1	14.3	5	1 UXAA_CHLTR	P38005 chlamydia t
91	1	14.3	5	1 ACPH_RABIT	P25154 oryctolagus
92	1	14.3	6	1 ASP2_LACSN	P325154 oryctolagus
93	1	14.3	6	1 C1P1_MYTED	P32655 lactobacill
94	1	14.3	6	1 C1P2_MYTED	P13736 mytilus edu
95	1	14.3	6	1 EI01_LITRU	P82096 litoria rub
96	1	14.3	6	1 FARP_MONEX	P41966 moniezia ex
97	1	14.3	6	1 LOK1_LOCM1	P41491 locusta mig
98	1	14.3	6	1 TMOF_SABRU	P41495 sarcophaga
99	1	14.3	6	1 TRP1_PSEPU	P36414 pseudomonas
100	1	14.3	6	1 UN06_CLOPA	P81351 clostridium
101	1	14.3	6	1 VP19_HSVIK	P23210 herpes simp
102	1	14.3	7	1 ALL2_CARMA	P81805 carcinus ma
103	1	14.3	7	1 ALL7_CYPVO	P82158 cydia pomon
104	1	14.3	7	1 BHPH_CONIM	P58803 conus imper
105	1	14.3	7	1 CARP_MYTED	P10420 mytilus edu
106	1	14.3	7	1 CCP1_ENTFA	P20104 enterococcu

107 1 14.3 7 1 CHO5_ALCSP
 108 1 14.3 7 1 E105_LITRU
 109 1 14.3 7 1 FAR1_ASCSU
 110 1 14.3 7 1 FAR1_HEITI
 111 1 14.3 7 1 FAR2_ASCSU
 112 1 14.3 7 1 FAR3_HAECO
 113 1 14.3 7 1 FAR3_PANRE
 114 1 14.3 7 1 FAR5_HIRME
 115 1 14.3 7 1 FARG_CALVO
 116 1 14.3 7 1 IGAO_DACDE
 117 1 14.3 7 1 LANC_CARDUI
 118 1 14.3 7 1 MNPI_LEPDE
 119 1 14.3 7 1 PH2_LYCES
 120 1 14.3 7 1 TV51_LITRU
 121 1 14.3 7 1 UC24_MAIZE
 122 1 14.3 7 1 UF03_MOUSE
 123 1 14.3 7 1 UH11_RAT
 124 1 14.3 7 1 UN06_PINPS
 125 1 14.3 7 1 WHA1_ACHFU
 126 1 14.3 7 1 WHA2_ACHFU
 127 1 14.3 7 1 WHA3_ACHFU
 128 1 14.3 8 1 ACT_THUAL
 129 1 14.3 8 1 ACT_CARMA
 130 1 14.3 8 1 AKH_LIBAU
 131 1 14.3 8 1 AKH_MELML
 132 1 14.3 8 1 AKH_TABAT
 133 1 14.3 8 1 AL12_CARMA
 134 1 14.3 8 1 AL17_CARMA
 135 1 14.3 8 1 AL18_CARMA
 136 1 14.3 8 1 ALL1_CYPDO
 137 1 14.3 8 1 ALL5_CYPDO
 138 1 14.3 8 1 ALL6_CYPDO
 139 1 14.3 8 1 ANG2_BOTJA
 140 1 14.3 8 1 CAD1_ENTFA
 141 1 14.3 8 1 CCKN_MACEU
 142 1 14.3 8 1 CLUP_THICU
 143 1 14.3 8 1 COM2_CONPU
 144 1 14.3 8 1 CPD1_ENTFA
 145 1 14.3 8 1 FAR1_PANRE
 146 1 14.3 8 1 FAR1_PENMO
 147 1 14.3 8 1 FAR2_MACRS
 148 1 14.3 8 1 FAR4_HOMAM
 149 1 14.3 8 1 FAR4_MACRS
 150 1 14.3 8 1 FAR7_ASCSU

ALIGNMENTS

RESULT 1
 ID TUFT HUMAN STANDARD; PRT; 4 AA.
 AC P01858;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Phagocytosis-stimulating peptide (Tuftsin).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RN SEQUENCE.

RP MEDLINE=72187087; PubMed=4112769;
 RA Nishioka K., Constantopoulos A., Satoh P.S., Najjar V.A.;
 RT "The characteristics, isolation and synthesis of the phagocytosis
 RT stimulating peptide tuftsin."
 RL Biochem. Biophys. Res. Commun. 47:172-179 (1972).
 RN [2]
 RP IMMUNOGLOBULIN CLASS.
 RP MEDLINE=68091045; PubMed=4169272;
 RA Fidalgo B.V., Najjar V.A.;
 RT "The physiological role of the lymphoid system. VI. The stimulatory

RT effect of leucophilic gamma globulin (leucokinin) on the phagocytic
 RT activity of human polymorphonuclear leucocyte.";
 RL Biochemistry 6:3386-3392(1967).
 CC -!- MISCELLANEOUS: AN IGG (CALLED LEUCOKININ) BINDS REVERSIBLY TO THE
 CC CELL MEMBRANE OF NEUTROPHILS IN THE BLOOD. LEUCOKININASE ON THE
 CC MEMBRANE RELEASES THE ACTIVE PEPTIDE TUFTSIN FROM THE GAMMA CHAIN.
 CC TUFTSIN IS ESSENTIAL FOR MAXIMUM STIMULATION OF THE PHAGOCYTIC
 CC ACTIVITY OF NEUTROPHILS.

DR PIR; A02147; A02147.
 DR MIM; 191150; -.
 DR GO; GO:0003823; F:antigen binding activity; NAS.
 DR GO; GO:0006909; P:phagocytosis; NAS.
 SQ SEQUENCE 4 AA; 501 WM; 74176321C0000000 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred.No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KP 5
 DB 2 KP 3

RESULT 2

OVN_LEPDE STANDARD; PRT; 6 AA.
 AC P42385;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Oviductal motility stimulating peptide (lep-OVM).
 OC Leptinotarsa decemlineata (Colorado potato beetle).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
 OC Phytophaga; Chrysomelidae; Chrysomelidae; Chrysomelinae;
 OC Chrysomelini; Leptinotarsa.
 OX NCBI_TaxID=7539;
 RN [1]
 RN SEQUENCE AND SYNTHESIS.

RP TISSUE=Head;
 RC MEDLINE=91271080; PubMed=2052497;
 RX Spittaels K., Schoofs L., Grauwels L., Smet H., van Damme J.,
 RA Proost P., Torrekens S., de Loof A.;
 RT "Isolation, identification and synthesis of novel oviductal motility
 RT stimulating head peptide in the Colorado potato beetle, Leptinotarsa
 RT decemlineata".
 RL Peptides 12:31-36(1991).
 CC -!- FUNCTION: MYOTROPIC PEPTIDE. STIMULATES THE CONTRACTIONS OF THE
 CC OVIDUCT.
 CC Neuropeptide; Amidation.
 FT MOD RES 6
 SQ SEQUENCE 6 AA; 720 WM; 6B07632B5DD03000 CRC64;

RESULT 3

ALL3_CARMA STANDARD; PRT; 7 AA.
 ID ALL3_CARMA
 AC P81806;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Carcinustatin 3.
 OS Carcinus maenas (Common shore crab) (Green crab).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;

OY 4 KP 5
 DB 4 KP 5

RESULT 3

ALL3_CARMA STANDARD; PRT; 7 AA.
 ID ALL3_CARMA
 AC P81806;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Carcinustatin 3.
 OS Carcinus maenas (Common shore crab) (Green crab).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;

OC Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX NCBI_TaxID=6759;

RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Multigene family.
SQ SEQUENCE 7 AA; 796 MW; 672879DCB476B70 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 PY 6
Db 2 PY 3

RESULT 4

ALL4_CARMA
ID -ALL4_CARMA STANDARD; PRT; 7 AA.

AC P81807;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 4.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX NCBI_TaxID=6759;

RN [1]
RP SEQUENCE
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Multigene family.
SQ SEQUENCE 7 AA; 782 MW; 672879DCB476AC0 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 PY 6
Db 2 PY 3

RESULT 5

ALL5_CARMA
ID -ALL5_CARMA STANDARD; PRT; 7 AA.

AC P81808;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 5.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidea; Portunidae; Carcinus.

OX NCBI_TaxID=6759;

RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Multigene family.
FT MOD RES 7 AMIDATION.
SQ SEQUENCE 7 AA; 781 MW; 672879DCB476420 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 PY 6
Db 2 PY 3

RESULT 6

PAR4_PANRE
ID -PAR4_PANRE STANDARD; PRT; 7 AA.

AC P41875;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE PMRFamide-like neuropeptide PF4 (KPNFIRP-amide).
OS Panagrellus redivivus.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
OC Panagrolaimoidea; Panagrolaimidae; Panagrellus.
OX NCBI_TaxID=6233;

RN [1]
RP SEQUENCE, AND SYNTHESIS.
RX MEDLINE=95232026; PubMed=7716079;
RA Maule A.G., Shaw C., Bowman J.W., Halton D.W., Thompson D.P.,
RA Thim L., Kubiak T.M., Martin R.A., Geary T.G.;
RT "Isolation and preliminary biological characterization of
RT KPNFIRFamide, a novel PMRFamide-related peptide from the free-living
RT nematode, Panagrellus redivivus.";
RL Peptides 16:87-93(1995).
CC -!- FUNCTION: MYOACTIVE; INDUCES A RAPID CONCENTRATION-DEPENDENT
CC -!- MUSCLE TENSION INCREASE.
CC -!- SIMILARITY: BELONGS TO THE FARP (PMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD RES 7 AMIDATION.

SQ SEQUENCE 7 AA; 921 MW; 69D40059C4576350 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KP 5
Db 1 KP 2

RESULT 7

GFRP_MOUSE
ID -GFRP_MOUSE STANDARD; PRT; 7 AA.

AC P99025;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE GTP cyclohydrolase I feedback regulatory protein (P35) (Fragment).
GN GCHFR OR GFRP.
OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Liver;
 RA Sanchez J.-C., Rouge V., Frutiger S., Hughes G., Yan J.X.,
 RA Hoogland C., Appel R.D., Binz P.-A., Hochstrasser D.F.,
 RA Cowthorne M.;
 RL Submitted (AUG-1998) to the SWISS-PROT data bank.
 CC -!- FUNCTION: MEDIATES TETRAHYDROBIOTERIN INHIBITION OF GTP
 CC CYCLOHYDROLASE I. THIS INHIBITION IS REVERSED BY L-PHENYLALANINE
 CC (BY SIMILARITY).
 CC -!- SUBUNIT: Homodimer (By similarity).
 DR SWISS-2DPAGE; P99025; MOUSE.
 FT INIT_MET 0
 FT NON_TER 7
 FT MOD_RES 3
 FT SEQUENCE 7 AA; 806 MW; 71B5B05273B4700 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PY 6
 ||
 Db 1 PY 2

RESULT 8
 ID HY7_PIG STANDARD; PRT; 7 AA.
 AC P01153;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 21-JUL-1986 (Rel. 01, Last annotation update)
 DE Hypothalamic heptapeptide.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE, AND SYNTHESIS.
 RX MEDLINE=81213980; PubMed=6263778;
 RA Chang R.C.-C., Huang W.-Y., Arimura A., Redding T.W., Coy D.H.,
 RA Saffran M., Kong A., Hamilton J.W., Cohn D.V., Schally A.V.;
 RT "Isolation, structure and synthesis of a heptapeptide with in vitro
 RT ACTH-releasing activity from porcine hypothalamus.";
 RL Horm. Metab. Res. 13:228-232(1981).
 DR PIR: A01417; NYPG7.
 SQ SEQUENCE 7 AA; 957 MW; 632B4581FB5059A0 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 YK 7
 ||
 Db 6 YK 7

RESULT 9
 ID TPFY_PACDA STANDARD; PRT; 7 AA.
 AC P83455;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Trypophyllin-1 (Pdt-1).
 OS Pachymedusa dactinolor (Giant mexican leaf frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
 OC Phyllomedusinae; Pachymedusa.
 OX NCBI_TaxID=75988;

[1]
 RN SEQUENCE, MASS SPECTROMETRY, HYDROXYLATION OF PRO-3, AND AMIDATION OF
 RP PRO-7.
 RC TISSUE=Skin secretion;
 RA Chen T.B., Orr D.F., Shaw C.;
 RA "Pachymedusa dactinolor tryptophyllin-1 (Pdt-1): structural
 RT characterization, pharmacological activity and cloning of precursor
 RT cDNA.";
 RL Submitted (SEP-2002) to the SWISS-PROT data bank.
 CC -!- FUNCTION: Myoactive. Has selective relaxing activity on vascular
 CC smooth muscle.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Skin.
 CC -!- MASS SPECTROMETRY: MW=809.2; METHOD=MALDI.
 DR GO: GO:0005576; C:extracellular; NAS.
 DR GO: GO:0045986; P:negative regulation of smooth muscle contra. . .; NAS.
 KW Amphibian defense peptide; Amidation; Hydroxylation.
 FT MOD_RES 3
 FT MOD_RES 7
 FT SEQUENCE 7 AA; 794 MW; 7772D37DC7776350 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KP 5
 ||
 Db 1 KP 2

RESULT 10
 ID UF04_MOUSE STANDARD; PRT; 7 AA.
 AC P18642;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE Unknown protein from 2D-page of fibroblasts (P46) (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=95009907; PubMed=7523108;
 RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
 RT "Separation and sequencing of familial and novel murine proteins
 RT using preparative two-dimensional gel electrophoresis.";
 RL Electrophoresis 15:735-745(1994).
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 5.0, ITS MW IS: 46 kDa.
 FT NON_TER 7
 FT SEQUENCE 7 AA; 766 MW; 68640AB777632700 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KP 5
 ||
 Db 2 KP 3

RESULT 11
 ID AKHG_GRYBI STANDARD; PRT; 8 AA.
 AC P14086;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Adipokinetic hormone G (AKH-G) (RO II).
 OX Gryllus bimaculatus (Two-spotted cricket), and

OS Romalea microptera (lubber grasshopper).
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
 OC Neoptera; Orthopteroidea; Orthoptera; Ensifera; Gryllidae; Gryllinae;
 OC Gryllus.
 OX NCBI_TaxID=6999, 7007;
 RN [1]
 RN SEQUENCE.
 RP SPECIES=G.bimaculatus; TISSUE=Corpora cardiaca;
 RX MEDLINE=88106553; PubMed=3426616;
 RA Gaede G., Rinehart K.L. Jr.;
 RT "Primary sequence analysis by fast atom bombardment mass spectrometry
 of a peptide with adipokinetic activity from the corpora cardiaca of
 the cricket Gryllus bimaculatus.";
 RL Biochem. Biophys. Res. Commun. 149:908-914(1987).
 RN [2]
 RN SEQUENCE.
 RP SPECIES=R.microptera; TISSUE=Corpora cardiaca;
 RX MEDLINE=89145002; PubMed=3226948;
 RA Gaede G., Hilbich C., Beyreuther K., Rinehart K.L. Jr.;
 RT "Sequence analyses of two neuropeptides of the AKH/RPCH-family from
 the lubber grasshopper, Romalea microptera.";
 RL Peptides 9:681-688(1988).
 CC -!- FUNCTION: THIS HORMONE, RELEASED FROM CELLS IN THE CORPORA
 CARDIACA AFTER THE BEGINNING OF FLIGHT, CAUSES RELEASE OF
 DIGLYCERIDES FROM THE FAT BODY AND THEN STIMULATES THE FLIGHT
 MUSCLES TO USE THESE DIGLYCERIDES AS AN ENERGY SOURCE.
 CC -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
 DR PIR: A28004; A28004.
 DR InterPro: IPR002047; AKH.
 DR PROSITE: PS00256; AKH; 1.
 KW Neuropeptide; Amidation; Flight; Pyrrolidone carboxylic acid.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 8 8 AMIDATION.
 SQ SEQUENCE 8 AA; 938 MW; 867861B5B9C452D6 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TG 2
 DB 6 TG 7

RESULT 12
 AL15 CARMA
 ID AL15 CARMA STANDARD; PRT; 8 AA.
 AC P81818;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Carcinustatin 15.
 OS Carcinus maenas (Common shore crab) (Green crab).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
 OC Eubrachyura; Portunioidea; Portunidae; Carcinus.
 OX NCBI_TaxID=6759;
 RN [1]
 RN SEQUENCE.
 RP TISSUE=Cerebral ganglion, and Thoracic ganglion;
 RX MEDLINE=98121193; PubMed=9461295;
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
 RT Thorpe A.;
 RT "Isolation and identification of multiple neuropeptides of the
 allatostatin superfamily in the shore crab Carcinus maenas.";
 RL Eur. J. Biochem. 250:727-734(1997).
 CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 KW Neuropeptide; Amidation; Multigene family.
 FT MOD_RES 8 8 AMIDATION.
 SQ SEQUENCE 8 AA; 811 MW; 922879D5AB47687D CRC64;

Query Match 28.6%; Score 2; DB 1; Length 8;

Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 5 PY 6
 DB 3 PY 4

RESULT 13
 AL16 CARMA
 ID AL16 CARMA STANDARD; PRT; 8 AA.
 AC P81819;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Carcinustatin 16.
 OS Carcinus maenas (Common shore crab) (Green crab).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
 OC Eubrachyura; Portunioidea; Portunidae; Carcinus.
 OX NCBI_TaxID=6759;
 RN [1]
 RN SEQUENCE.
 RP TISSUE=Cerebral ganglion, and Thoracic ganglion;
 RX MEDLINE=98121193; PubMed=9461295;
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
 RT Thorpe A.;
 RT "Isolation and identification of multiple neuropeptides of the
 allatostatin superfamily in the shore crab Carcinus maenas.";
 RL Eur. J. Biochem. 250:727-734(1997).
 CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 KW Neuropeptide; Amidation; Multigene family.
 FT MOD_RES 8 8 AMIDATION.
 SQ SEQUENCE 8 AA; 813 MW; 7C286845AB476878 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 PY 6
 DB 3 PY 4

RESULT 14
 ALL3 CYDPO
 ID ALL3 CYDPO STANDARD; PRT; 8 AA.
 AC P82154;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Cydiastatin 3.
 OS Cydia pomonella (Codling moth).
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Tortricoidae; Tortricidae; Olethreutinae; Cydia.
 OX NCBI_TaxID=82600;
 RN [1]
 RN SEQUENCE.
 RP TISSUE=Larva;
 RX MEDLINE=98054539; PubMed=9392829;
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D.,
 RT Davey M., East P.D., Thorpe A.;
 RT "Lepidopteran peptides of the allatostatin superfamily.";
 RL Peptides 18:1301-1309(1997).
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 KW Neuropeptide; Amidation.
 FT MOD_RES 8 8 AMIDATION.
 SQ SEQUENCE 8 AA; 926 MW; C82879D5AB477415 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PY 6
 DB ||
 3 PY 4

RESULT 15

ALL4_CVDPO
 ID ALL4_CVDPO STANDARD; PRT; 8 AA.
 AC P41840;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Callatostatin 4 (Leu-callatostatin 4).
 OS Calliphora vomitoria (Blue blowfly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
 OC Calliphoridae; Calliphora.
 OX NCBI_TaxID=27454;
 RN [1]
 RP SEQUENCE.

TISSUE=Thoracic ganglion;
 RX MEDLINE=932111980; PubMed=8460157;
 RA Duve H., Johnsen A.H., Scott A.G., Yu C.G., Yagi K.J., Tobe S.S.,
 RA Thorpe A.;
 RT "Callatostatin: neuropeptides from the blowfly Calliphora vomitoria
 RT with sequence homology to cockroach allatostatins.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:2456-2460(1993).
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=94291167; PubMed=8020069;
 RA Duve H., Thorpe A.;
 RT "Distribution and functional significance of Leu-callatostatins in
 RT the blowfly Calliphora vomitoria.";
 RL Cell Tissue Res. 276:367-379(1994).

CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR AND PLAY
 CC A ROLE IN THE INTEGRATION OF INFORMATION WITHIN THE BRAIN. MAY BE
 CC INVOLVED IN THE CONTROL OF VISCERAL MUSCLES DUE TO ITS ABILITY TO
 CC BEHAVE AS POTENT INHIBITORS OF PERISTALTIC MOVEMENTS. MAY ALSO
 CC FULFILL A NEUROHORMONAL ROLE ON MUSCLES OF THE GUT AND HEART.
 CC -!- TISSUE SPECIFICITY: BRAIN, SUBESOPHAGEAL GANGLION, RETROCEBRAL
 CC COMPLEX, THORACICO-ABDOMINAL GANGLION, PERIPHERAL NEUROSECRETORY
 CC SYSTEM AND INTESTINE.
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.

DR PIR; D47393;
 FT MOD RES 8 1
 FT UNSURE 8 1
 SQ SEQUENCE 8 AA; 954 MW; D32879D5AB47740A CRC64;
 OR N.
 AMIDATION.
 Query Match 28.6%; Score 2; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PY 6
 DB ||
 3 PY 4

RESULT 16

ALL4_CVDPO
 ID ALL4_CVDPO STANDARD; PRT; 8 AA.
 AC P82155;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE Callatostatin 4.
 OS Cydia pomonella (Coddling moth).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Tortricidae; Tortricidae; Olethreutinae; Cydia.
 OX NCBI_TaxID=82600;

RN [1]
 RP SEQUENCE.
 RC TISSUE=Larva;
 RX MEDLINE=98054539; PubMed=9392829;
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D.,
 RA Davey M., East P.D., Thorpe A.;
 RT "Lepidopteran peptides of the allatostatin superfamily.";
 RL Peptides 18:1301-1309(1997).
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 KW Neuropeptide; Amidation.
 FT MOD RES 8 8
 SQ SEQUENCE 8 AA; 910 MW; 922879D5AB47740D CRC64;
 AMIDATION.
 Query Match 28.6%; Score 2; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PY 6
 DB ||
 3 PY 4

RESULT 17

ALL5_CALVO
 ID ALL5_CALVO STANDARD; PRT; 8 AA.
 AC P41841;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Callatostatin 5 (Met-callatostatin 1) ([Hyp3]Met-callatostatin).
 OS Calliphora vomitoria (Blue blowfly).
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
 OC Calliphoridae; Calliphora.
 OX NCBI_TaxID=27454;
 RN [1]

RN SEQUENCE.
 RP MEDLINE=932111980; PubMed=8460157;
 RA Duve H., Johnsen A.H., Scott A.G., Yu C.G., Yagi K.J., Tobe S.S.,
 RA Thorpe A.;
 RT "Callatostatin: neuropeptides from the blowfly Calliphora vomitoria
 RT with sequence homology to cockroach allatostatins.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:2456-2460(1993).
 RN [2]

RP CHARACTERIZATION, AND HYDROXYLATION.
 RC TISSUE=Head;
 RX MEDLINE=94342269; PubMed=8063725;
 RA Duve H., Johnsen A.H., Scott A.G., East P., Thorpe A.;
 RT "[Hyp3]Met-callatostatin. Identification and biological properties of
 RT a novel neuropeptide from the blowfly Calliphora vomitoria.";
 RL J. Biol. Chem. 269:21059-21066(1994).

CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR AND PLAY
 CC A ROLE IN THE INTEGRATION OF INFORMATION WITHIN THE BRAIN. MAY BE
 CC INVOLVED IN THE CONTROL OF VISCERAL MUSCLES DUE TO ITS ABILITY TO
 CC BEHAVE AS POTENT INHIBITORS OF PERISTALTIC MOVEMENTS. MAY ALSO
 CC FULFILL A NEUROHORMONAL ROLE ON MUSCLES OF THE GUT AND HEART.
 CC -!- TISSUE SPECIFICITY: NEURONS WITHIN BRAIN AND ABDOMINAL GANGLION.

CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 DR PIR; E47393;
 KW Neuropeptide; Amidation; Hydroxylation.
 FT MOD RES 3 3
 FT UNSURE 8 8
 SQ SEQUENCE 8 AA; 883 MW; 7D9879CABB477768 CRC64;
 AMIDATION.

Query Match 28.6%; Score 2; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PY 6
 DB ||
 3 PY 4

RESULT 18

ALL7_CARMA
ID ALL7_CARMA STANDARD; PRT; 8 AA.
AC P81809; P81804; P81810;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Carcinustatin 7 [Contains: Carcinustatin 6; Carcinustatin 1].
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation; Multigene family.
FT PEPTIDE 1 8 CARCINUSTATIN 7.
FT PEPTIDE 2 8 CARCINUSTATIN 6.
FT PEPTIDE 4 8 CARCINUSTATIN 1.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 825 MW; 922879CDB4775BD CRC64;

Query Match 28.6%; Score 2; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 PY 6
Db 3 PY 4

RESULT 19

ALL8_CARMA
ID ALL8_CARMA STANDARD; PRT; 8 AA.
AC P81811;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 8.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation; Multigene family.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 795 MW; 922879CDB47687D CRC64;

Query Match 28.6%; Score 2; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 PY 6

Db 3 PY 4

RESULT 20

ALL9_CARMA
ID ALL9_CARMA STANDARD; PRT; 8 AA.
AC P81812;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 9.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation; Multigene family.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 781 MW; 7C2879CDB476878 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 PY 6
Db 3 PY 4

RESULT 21

B44K_FORGI
ID B44K_FORGI STANDARD; PRT; 8 AA.
AC P81886;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 44 kDa immunogenic protein (Fragment).
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP SEQUENCE.
RC STRAIN=VPR 3492;
RX MEDLINE=20198497; PubMed=10731616;
RA Norris J.M., Love D.N.;
RT "Serum antibody responses of cats to soluble whole cell antigens of
feline Porphyromonas gingivalis.";
RL Vet. Microbiol. 73:37-49(2000).
CC -!- SIMILARITY: TO P.GINGIVALIS HEMAGGLUTININ A.
KW Antigen.
FT NON_TER
SQ SEQUENCE 8 AA; 989 MW; 9554540326CB476D CRC64;

Query Match 28.6%; Score 2; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 PY 6
Db 2 PY 3

```

RESULT 22
ID ORMY ORCLI STANDARD; PRT; 8 AA.
AC P82455;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Orcomytropin (OMT).
OS Orconectes limosus (Spinycheek crayfish).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
OC Astacodea; Cambaridae; Orconectes.
OX NCBI_TaxID=28379;
RN [1]
SEQUENCE, MASS SPECTROMETRY, AND AMIDATION.
RP TISSUE=Hindgut;
RC MEDLINE=20411310; PubMed=10952880;
RA Dirksen H., Burdick S., Sauter A., Keller R.;
RT "Two orokinins and the novel octapeptide orcomytropin in the hindgut
of the crayfish Orconectes limosus: identified myostimulatory
neuropeptides originating together in neurones of the terminal
abdominal ganglion."
RL J. Exp. Biol. 203:2807-2818(2000).
CC -!- FUNCTION: MYOTROPIC PEPTIDE, ENHANCES BOTH THE FREQUENCY AND
AMPLITUDE OF SPONTANEOUS HINDGUT CONTRACTIONS. IT IS SYNTHESIZED
BY ABDOMINAL GANGLIONIC NEURONS
CC -!- MASS SPECTROMETRY: MW=904.8; METHOD=FAB.
KW Amidation; Neuropeptide.
FT MOD RES 8
SQ SEQUENCE 8 AA; 905 MW; 87C861B1A9CDDAA9 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
DB 6 TG 7

RESULT 23
RS1 ERWCH STANDARD; PRT; 8 AA.
AC P37585;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE 30S ribosomal protein S1 (fragment).
GN RPSA.
OS Erwinia chrysanthemi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=556;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=3937;
RA Doullie A., Toussaint A., Faelen M.;
RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: BINDS MRNA; THUS FACILITATING RECOGNITION OF THE
SHINE-DALGARNO (SD) PURINE-RICH SEQUENCE (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE S1P FAMILY OF RIBOSOMAL PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC

RESULT 24
AL10_CARMA STANDARD; PRT; 9 AA.
ID AL10_CARMA
AC P81813;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 10.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
SEQUENCE.
RP TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
allatostatin superfamily in the shore crab Carcinus maenas."
RL Eur. J. Biochem. 250:727-734(1997).
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation; Multigene family.
FT MOD RES 9
SQ SEQUENCE 9 AA; 963 MW; 372D79CDBC4776C7 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PY 6
DB 4 PY 5

RESULT 25
AL11_CARMA STANDARD; PRT; 9 AA.
ID AL11_CARMA
AC P81814;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 11.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
SEQUENCE.
RP TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
allatostatin superfamily in the shore crab Carcinus maenas."

```

```

RL Bur. J. Biochem. 250:727-734(1997).
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation; Multigene family.
FT MOD_RES 9 AMIDATION.
SQ SEQUENCE 9 AA; 927 MW; 832D79CDB46D861 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TG 2
Db 2 TG 3

RESULT 26
BUK_CLOPA STANDARD; PRT; 9 AA.
AC P81337;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Butyrate kinase (EC 2.7.2.7) (BK) (CP 38) (Fragment).
GN BUK.
OS Clostridium pasteurianum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1501;
RN [1]
RP SEQUENCE.
RC STRAIN=W5;
RA Flegstrud R., Skjeldal L.;
RX MEDLINE=98291870; PubMed=9629918;
RT "Two-dimensional gel electrophoresis separation and N-terminal
sequence analysis of proteins from Clostridium pasteurianum W5.";
RL Electrophoresis 19:802-806(1998).
CC -!- FUNCTION: Catalyzes the conversion of butyryl-CoA through butyryl
phosphate to butyrate (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + 2-butanolate = ADP + butanoyl phosphate.
CC -!- PATHWAY: Involved in the acidogenic phase of fermentation.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the acetokinase family.
DR HAMAP; MF 00542; -; 1.
DR InterPro; IPR000890; Acetate kin.
DR PROSITE; PS01075; ACETATE KINASE 1; PARTIAL.
DR PROSITE; PS01076; ACETATE_KINASE_2; PARTIAL.
KW Transferase; Kinase.
FT NON_TER 9
SQ SEQUENCE 9 AA; 1104 MW; 055D40505727232B CRC64;

Query Match 28.6%; Score 2; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 YK 7
Db 2 YK 3

RESULT 27
CCAP_CARMA STANDARD; PRT; 9 AA.
AC P38556;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Cardioactive peptide (CCAP).
OS Carcinus maenas (Common shore crab) (Green crab),
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm),
OS Tenebrio molitor (Yellow mealworm), and
OS Spodoptera eridania (Southern armyworm).
OS Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;

```

```

OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Brachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_TaxID=6759, 7130, 7067, 37547;
RN [1]
RP SEQUENCE.
RC SPECIES=C.maenas; TISSUE=Pericardial organs;
RA Stangier J., Hilbich C., Beyreuther K., Keller R.;
RT "Unusual cardioactive peptide (CCAP) from pericardial organs of the
shore crab Carcinus maenas";
RL Proc. Natl. Acad. Sci. U.S.A. 84:575-579(1987).
RN [2]
RP SEQUENCE.
RC SPECIES=M.sexata;
RX MEDLINE=93050243; PubMed=1426284;
RA Cheung C.C., Loi P.K., Sylwester A.W., Lee T.D., Tublitz N.J.;
RT "Primary structure of a cardioactive neuropeptide from the tobacco
hawkmoth, Manduca sexta.";
RL FEBS Lett. 313:165-168(1992).
RN [3]
RP SEQUENCE.
RC SPECIES=T.molitor, and S.eridania; TISSUE=Head;
RX MEDLINE=94176032; PubMed=8129851;
RA Furuya K., Liao S., Reynolds S.E., Ota R.B., Hackett M.,
RA Schooley D.A.;
RT "Isolation and identification of a cardioactive peptide from Tenebrio
molitor and Spodoptera eridania.";
RL Biol. Chem. Hoppe-Seyler 374:1065-1074(1993).
CC -!- FUNCTION: THE EFFECT OF CCAP IS BOTH INO- AND CHRONOTROPIC.
CC -!- TISSUE SPECIFICITY: STORED IN PERICARDIAL ORGANS AND RELEASED
INTO THE HEMOLYPH.
DR PIR; A26363; A26363.
DR PIR; S27233; S27233.
KW Neuropeptide; Amidation.
FT DISULFID 3 9
MOD_RES 9 AA; 959 MW; CSA861A9CDD44EB9 CRC64;
SQ SEQUENCE 9 AA; 959 MW; CSA861A9CDD44EB9 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TG 2
Db 7 TG 8

RESULT 28
COW_CONVE STANDARD; PRT; 9 AA.
AC P83047;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Contryphan-Vn.
OS Conus ventricosus (Mediterranean cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=117992;
RN [1]
RP SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.
RC TISSUE=Venom;
RX MEDLINE=21547785; PubMed=11688995;
RA Masbilia G.R., Schinina M.E., Ascenzi P., Politicelli F.;
RT "Contryphan-Vn: a novel peptide from the venom of the Mediterranean
snail Conus ventricosus";
RL Biochem. Biophys. Res. Commun. 288:908-913(2001).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -!- MASS SPECTROMETRY: MW=1088.6; METHOD=MALDI.
CC -!- SIMILARITY: BELONGS TO THE CONTRYPHAN FAMILY.
KW Toxin; Amidation; D-amino acid.
FT DISULFID 3 9

```

FT MOD_RES 5 5 D-TRYPTOPHAN.
 FT MOD_RES 9 9 AMIDATION
 SQ SEQUENCE 9 AA; 1091 MW; 8D38676323676BBA CRC64;
 Query Match 28.6%; Score 2; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KP 5
 DB 6 KP 7

RESULT 29
 COXE_THUOB
 ID COXE_THUOB STANDARD; PRT; 9 AA.
 AC P80975;
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Cytochrome c oxidase polypeptide Via (EC 1.9.3.1) (Fragment).
 OS Thunus obesus (Bigeye tuna).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
 OC Scombridae; Thunnus.
 OX NCBI_TaxID=8241;
 RN [1]

SEQUENCE.
 RP TISSUE=Heart;
 RC MEDLINE=97454291; PubMed=9310366;
 RA Arnold S., Lee I., Kim M., Song E., Linder D., Lottspeich F.,
 RA Kadenbach B.;
 RT "The subunit structure of cytochrome-c oxidase from tuna heart and
 RT liver.";
 RL Eur. J. Biochem. 248:99-103(1997).
 CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
 CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
 CC MITOCHONDRIAL ELECTRON TRANSPORT.
 CC -!- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
 CC C + 2 H(2)O.
 CC -!- SUBCELLULAR LOCATION: Mitochondrial inner membrane.
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIA FAMILY.
 DR PIR; S77984; S77984.
 DR InterPro: IPR001349; COX6A.
 DR PROSITE; PS01329; COX6A; PARTIAL.
 KW Oxidoreductase; Inner membrane; Mitochondrion.
 FT NON_TER 1 1
 FT NON_TER 9 9
 SQ SEQUENCE 9 AA; 1136 MW; 62E072C9CB0776DB CRC64;
 Query Match 28.6%; Score 2; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PY 6
 DB 8 PY 9

RESULT 30
 DSIP_RABIT
 ID DSIP_RABIT STANDARD; PRT; 9 AA.
 AC P01158;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Delta sleep-inducing peptide (DSIP).
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]

RP SEQUENCE.
 RX MEDLINE=77185324; PubMed=862769;
 RA Monnier M., Dudler L., Gachter R., Maier P.F., Tobler H.J.,
 RA Schoenenberger G.A.;
 RT "The delta sleep inducing peptide (DSIP). Comparative properties of
 RT the original and synthetic nonapeptide.";
 RL Experientia 33:548-552(1977).
 RN [2]

RP SEQUENCE, AND SYNTHESIS.
 RX MEDLINE=79054421; PubMed=568769;
 RA Schoenenberger G.A., Maier P.F., Tobler H.J., Wilson K., Monnier M.;
 RT "The delta EEG (sleep)-inducing peptide (DSIP). XI. Amino-acid
 RT analysis, sequence, synthesis and activity of the nonapeptide.";
 RL Pflugers Arch. 376:119-129(1978).
 RN [3]

REVIEW.
 RP MEDLINE=87175129; PubMed=3550726;
 RX Graf M.V., Kastin A.J.;
 RA "Delta-sleep-inducing peptide (DSIP): an update.";
 RL Peptides 7:1165-1187(1986).
 CC -!- FUNCTION: WHEN INFUSED INTO THE MESODIENCEPHALIC VENTRICLE OF
 CC RECIPIENT RABBITS INDUCES SPINDLE AND DELTA EEG ACTIVITY AND
 CC REDUCED MOTOR ACTIVITIES.
 CC -!- MISCELLANEOUS: THIS PEPTIDE WAS OBTAINED FROM DIALYSATES OF
 CC OCCIPITAL VENOUS SINUS BLOOD FROM RABBITS KEPT ASLEEP BY ELECTRIC
 CC STIMULATION OF THE THALAMUS.
 CC -!- DATABASE: NAME=Protein Spotlight;
 CC NOTE=issue 8 of March 2001;
 CC PIR; A01422; QDRB.
 DR WWW="http://www.expasy.org/spotlight/articles/sptlt008.html".
 SQ SEQUENCE 9 AA; 849 MW; DDD365BDDAA8787D CRC64;

Query Match 28.6%; Score 2; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GE 3
 DB 8 GE 9

RESULT 31
 FAR4_CALVO
 ID FAR4_CALVO STANDARD; PRT; 9 AA.
 AC P41859;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE CalliFMRFamide 4.
 OS Calliphora vomitoria (Blue blowfly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
 OC Calliphoridae; Calliphora.
 OX NCBI_TaxID=27454;
 RN [1]

SEQUENCE.
 RP TISSUE=Thoracic ganglion;
 RC MEDLINE=92196111; PubMed=1549595;
 RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,
 RA Rehfeld J.F., Thorpe A.;
 RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
 RT neuropeptides (designated calliFMRFamides) from the blowfly
 RT Calliphora vomitoria.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
 CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
 CC FAMILY.
 DR PIR; D41978; D41978.
 KW Neuropeptide; Amidation.
 FT MOD_RES 9
 SQ SEQUENCE 9 AA; 1182 MW; 31730699CAB6D457 CRC64;
 Query Match 28.6%; Score 2; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KP 5
 DB 1 KP 2

RESULT 32
 FAR5_ASCSU
 ID FAR5_ASCSU STANDARD; PRT; 9 AA.
 AC P43170;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DE FMRPamide-like neuropeptide AF5.
 DE FMRPamide-like neuropeptide AF5.
 OS Ascaris suum (Pig roundworm) (Ascaris lumbricoides).
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
 OC Ascarididae; Ascaris.
 OX NCBI_TaxID=6253;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=95380362; PubMed=7651904;
 RA Cowden C., Stretton A.O.W.;
 RA "Eight novel FMRPamide-like neuropeptides isolated from the nematode
 RT Ascaris suum.";
 RL Peptides 16:491-500(1995).
 CC -!- SIMILARITY: BELONGS TO THE FARP (FMRPAMIDE RELATED PEPTIDE)
 CC FAMILY.
 KW Neuropeptide; Amidation.
 FT MOD_RES 9 9
 SQ SEQUENCE 9 AA; 1052 MW; 340B0059D1B76338 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KP 5
 DB 3 KP 4

RESULT 33
 FLA2_TREHY
 ID FLA2_TREHY STANDARD; PRT; 9 AA.
 AC P80159;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Flagellar filament outer layer protein flaA2 (35 kDa sheath protein)
 DE (Fragment).
 GN FLA2.
 OS Treponema hyodysenteriae (Serpulina hyodysenteriae).
 OC Bacteria; Spirochaetes; Spirochaetales; Brachyspiraceae; Brachyspira.
 OX NCBI_TaxID=159;
 RN [1]
 RP SEQUENCE.
 RX STRAIN=C5;
 RX MEDLINE=93139764; PubMed=1487733;
 RA Koopman M.B.H., Baats E., van Vorstenbosch C.J.A.H.V.,
 RA van der Zeijst B.A.M., Kusters J.G.;
 RT "The periplasmic flagella of Serpulina (Treponema) hyodysenteriae are
 RT composed of two sheath proteins and three core proteins.";
 RL J. Gen. Microbiol. 138:2697-2706(1992).
 CC -!- FUNCTION: COMPONENT OF THE OUTER LAYER OF THE FLAGELLA.
 CC -!- SUBUNIT: THE FLAGELLUM CONSISTS OF AN OUTER LAYER COMPOSED OF TWO
 CC SHEATH PROTEINS, FLA1 (44 kDa) AND FLA2 (35 kDa) AROUND A CORE
 CC THAT CONTAINS THREE PROTEINS FLA3 (37 kDa), FLA2 (34 kDa) AND
 CC FLA3 (32 kDa).
 CC -!- SUBCELLULAR LOCATION: Periplasmic flagellum.
 KW Flagella; Periplasmic.
 FT UNSURE 2 2
 FT UNSURE 8 9
 FT NON_TER 9 9

SQ SEQUENCE 9 AA; 1129 MW; 855A19C68B4772D1 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PY 6
 DB 4 PY 5

RESULT 34
 LITR_PHYRO
 ID LITR_PHYRO STANDARD; PRT; 9 AA.
 AC P08946;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Rhodel-litorin.
 OS Phyllomedusa rohdei (Rohde's leaf frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
 OC Phyllomedusinae; Phyllomedusa.
 OX NCBI_TaxID=8394;
 RN [1]
 RP SEQUENCE.
 RX TISSUE=Skin secretion;
 RX MEDLINE=85127560; PubMed=3838283;
 RA Barra D., Byspamer G.F., Simmaco M., Bossa F., Melchiorri P.,
 RA Byspamer V.;
 RT "Rohdei-litorin: a new peptide from the skin of Phyllomedusa rohdei.";
 RL FEBS Lett. 182:53-56(1985).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Skin.
 CC -!- SIMILARITY: BELONGS TO THE BOMBESIN/NEUROMEDIN B/RANATENSIN
 CC FAMILY.
 DR PIR; S07241; S07241.
 DR InterPro; IPR000874; Bombesin.
 DR Pfam; PF02044; Bombesin; 1.
 DR PROSITE; PS00257; BOMBESIN; 1.
 KW Amphibian defense peptide; Bombesin family; Amidation;
 KW Pyrrolidone carboxylic acid.
 FT MOD_RES 1 1
 FT MOD_RES 9 9
 SQ SEQUENCE 9 AA; 1090 MW; 4ECCC1B861ADC377 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
 DB 5 TG 6

RESULT 35
 NEUX_HUMAN
 ID NEUX_HUMAN STANDARD; PRT; 9 AA.
 AC P04277;
 DT 20-MAR-1987 (Rel. 04, Created)
 DT 20-MAR-1987 (Rel. 04, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Neurotensin-related peptide (NRP) (Kinetensin).
 OS Homo sapiens (human).
 OS Bos taurus (Bovine), and
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9605, 9913, 9986;
 RN [1]
 RP SEQUENCE.
 RX SPECIES=Human;
 RX MEDLINE=86242180; PubMed=3087352;

RA Mogard M.H., Kobayashi R., Chen C.F., Lee T.D., Reeve J.R. Jr.,
 RA Shively J.E., Walsh J.H.;
 RT "The amino acid sequence of kinetensin, a novel peptide isolated from
 RT pepsin-treated human plasma: homology with human serum albumin,
 RT neurotensin and angiotensin.";
 RL Biochem. Biophys. Res. Commun. 136:983-988(1986).
 RP [2]
 RP SEQUENCE.
 RC SPECIES=Human, Bovine, and Rabbit;
 RX MEDLINE=87194805; PubMed=2437111;
 RA Caraway R.E., Mitra S.P., Cochrane D.E.;
 RT "Structure of a biologically active neurotensin-related peptide
 RT obtained from pepsin-treated albumin(s).";
 RL J. Biol. Chem. 282:5968-5973(1987).
 CC -I- FUNCTION: REGULATION OF FAT DIGESTION, LIPID ABSORPTION, AND
 CC BLOOD FLOW (POTENTIAL).
 CC PIR; A38885; ABBOS.
 DR GO; GO:0005576; C:extracellular; NAS.
 DR GO; GO:0005180; F:peptide hormone; NAS.
 DR GO; GO:0007599; P:hemosstasis; IDA.
 DR GO; GO:0006629; P:lipid metabolism; NAS.
 KW Hormone.
 SQ SEQUENCE 9 AA; 1172 MW; C804DB4761F4140D CRC64;
 Query Match 28.6%; Score 2; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 PY 6
 DB 6 PY 7
 RESULT 36
 ID OXYT EISFO STANDARD; PRT; 9 AA.
 AC P42998;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Annetocin.
 OS Eisenia foetida (Common branding worm) (Common dung-worm).
 OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
 OC Lumbricina; Lumbricidae; Eisenia.
 OX NCBI_TaxID=6396;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Pituitary;
 RX MEDLINE=94121660; PubMed=8292046;
 RA Oumi T., Ukena K., Matsushima O., Ikeda T., Fujita T., Minakata H.,
 RA Annetocin: an oxytocin-related peptide isolated from the earthworm,
 RT Eisenia foetida.";
 RL Biochem. Biophys. Res. Commun. 198:393-399(1994).
 CC -I- FUNCTION: POTENTIATES SPONTANEOUS CONTRACTIONS OF THE GUT AND ALSO
 CC PULSATORY CONTRACTIONS AND BLADDER-SHAKING MOVEMENT OF THE
 CC NEPHRIDIA. MAY BE INVOLVED IN OSMOREGULATION OF THE ANIMAL THROUGH
 CC NEPHRIDIAL FUNCTION.
 CC -I- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
 DR PIR; PC2021; PC2021.
 DR InterPro; IPR000981; Neurhyp_horm.
 DR Pfam; PF00220; hormone4; 1.
 DR PROSITE; PS00264; NEUROHYPOPHYS_HORM; FALSE_NEG.
 KW Hormone; Amidation.
 FT DISULFID 1 6
 FT MOD_RES 9 9 AMIDATION.
 SQ SEQUENCE 9 AA; 996 MW; D4EEB76EB45412C9 CRC64;
 Query Match 28.6%; Score 2; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TG 2

DB 8 TG 9
 RESULT 37
 ID ULAE HUMAN STANDARD; PRT; 9 AA.
 AC P31931;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Unknown protein from 2D-page of liver tissue (Spot 115) (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Liver;
 RX MEDLINE=94147969; PubMed=8313870;
 RA Hughes G.J., Frutiger S., Paquet N., Pasquali C., Sanchez J.-C.,
 RA Tissot J.-D., Bairoch A., Appel R.D., Hochstrasser D.F.;
 RT "Human liver protein map: update 1993.";
 RL Electrophoresis 14:1216-1222(1993).
 CC -I- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 5.5, ITS MW IS: 11 kDa.
 DR SWISS-2DPAGE; P31931; HUMAN.
 FT NON_TER 9
 SQ SEQUENCE 9 AA; 1096 MW; 477B3B173AE729C7 CRC64;
 Query Match 28.6%; Score 2; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 EX 4
 DB 7 EX 8
 RESULT 38
 ID UPAT7 HUMAN STANDARD; PRT; 9 AA.
 AC P30093;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Unknown protein from 2D-page of plasma (Spot 18) (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Plasma;
 RX MEDLINE=93092937; PubMed=1459097;
 RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,
 RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,
 RA Hochstrasser D.F.;
 RT "Plasma protein map: an update by microsequencing.";
 RL Electrophoresis 13:707-714(1992).
 CC -I- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 5.05, ITS MW IS: 37 kDa.
 DR SWISS-2DPAGE; P30093; HUMAN.
 FT NON_TER 1 1
 FT UNSURE 5 5
 FT NON_TER 9 9
 SQ SEQUENCE 9 AA; 1042 MW; 5C14477AE80772C7 CRC64;
 Query Match 28.6%; Score 2; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 PY 6

```

Db      ||
       7 PY 8

RESULT 39
CA12 LITCI
ID CA12 LITCI STANDARD; PRT; 10 AA.
AC P82086;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Caerulein 1.2/1.2Y4.
OS Litoria citropa (Australian blue mountains tree frog), and
OS Litoria splendida (Magnificent tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Pelodyadinae; Litoria.
OX NCBI_TaxID=94770, 30345;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY (CAERULEINS 1.2 AND 1.2Y4).
RC SPECIES=L.citropa; TISSUE=Skin secretion;
RX MEDLINE=20057701; PubMed=10589099;
RA Wabnitz P.A., Bowie J.H., Tyler M.J.;
RT "Caerulein-like peptides from the skin glands of the Australian blue
RT mountains tree frog Litoria citropa. Part 1. Sequence determination
RT using electrospray mass spectrometry."
RL Rapid Commun. Mass Spectrom. 13:2498-2502 (1999).
RN [2]
RP SEQUENCE, AND MASS SPECTROMETRY (CAERULEIN 1.2).
RC SPECIES=L.splendida; TISSUE=Skin secretion;
RX MEDLINE=20069371; PubMed=10601876;
RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C., Smith B.P.;
RT "Differences in the skin peptides of the male and female Australian
RT tree frog Litoria splendida. The discovery of the aquatic male sex
RT pheromone splendipherin, together with Phe8 caerulein and the
RT antibiotic peptide caerin 1.10."
RL Eur. J. Biochem. 267:269-275 (2000).
CC -!- FUNCTION: HYPOTENSIVE NEUROPEPTIDE (PROBABLE).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Skin dorsal glands.
CC -!- PTM: Isoform 1.2Y4 differs from isoform 1.2 in not being
CC sulfated.
CC -!- MASS SPECTROMETRY: MW=1366; METHOD=Electrospray.
CC -!- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.
DR InterPro; IPR001651; Gastrin.
DR PROSITE; PS00259; GASTRIN; FALSE NEG.
KW Amphibian defense peptide: Hypotensive agent; Amidation; Sulfation;
KW Pyrrolidone carboxylic acid. PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 1 1
FT MOD_RES 4 4
FT MOD_RES 10 10
SQ SEQUENCE 10 AA; 1306 MW; 99DBFCD37861BB5A CRC64;

Query Match 28.6%; Score 2; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 6.9e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
   ||
Db 5 TG 6

RESULT 40
CAER LITXA
ID CAER LITXA STANDARD; PRT; 10 AA.
AC P56264;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Caerulein.
OS Litoria xanthomera (Orange-thighed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;

```

```

OC Pelodyadinae; Litoria.
OX NCBI_TaxID=79697;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE=Skin secretion;
RX MEDLINE=97374000; PubMed=9230483;
RA Steinbörner S.T., Waugh R.J., Bowie J.H., Wallace J.C., Tyler M.J.,
RA Ramsay S.L.;
RT "New caerin antibacterial peptides from the skin glands of the
RT Australian tree frog Litoria xanthomera."
RL J. Pept. Sci. 3:181-185 (1997).
CC -!- FUNCTION: HYPOTENSIVE NEUROPEPTIDE.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Skin dorsal glands.
CC -!- MASS SPECTROMETRY: MW=1354; METHOD=FAB.
CC -!- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.
DR InterPro; IPR001651; Gastrin.
DR PROSITE; PS00259; GASTRIN; 1.
KW Amphibian defense peptide; Hypotensive agent; Amidation; Sulfation;
KW Pyrrolidone carboxylic acid. PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 1 1
FT MOD_RES 4 4
FT MOD_RES 10 10
SQ SEQUENCE 10 AA; 1290 MW; 99DBF3837861BB5A CRC64;

Query Match 28.6%; Score 2; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 6.9e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
   ||
Db 5 TG 6

RESULT 41
COXQ RABIT
ID COXQ RABIT STANDARD; PRT; 10 AA.
AC P80336;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome c oxidase polypeptide VIII-liver/heart (EC 1.9.3.1)
DE (Fragment).
GN COX8H.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE.
RC TISSUE=Heart, and Liver;
RC TISSUE=Heart, and Liver;
RA Freund R., Kadenbach B.;
RL Submitted (MAR-1994) to the SWISS-PROT data bank.
CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC MITOCHONDRIAL ELECTRON TRANSPORT.
CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIII FAMILY.
KW Oxidoreductase; Mitochondrion.
FT NON_TER 10
FT NON_TER 10
SQ SEQUENCE 10 AA; 1027 MW; 2C325CB40DC76338 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 6.9e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KP 5
   ||
Db 4 KP 5

RESULT 42

```

COXQ SHEEP
ID COXQ SHEEP STANDARD; PRT; 10 AA.
AC P80337; 1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome c oxidase polypeptide VIII-liver/heart (EC 1.9.3.1)
DE (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
ON NCBI_TaxID=9940;
RN [1]
RP SEQUENCE.
RC TISSUE=Heart, and Liver;
RA Freund R., Kadenbach B.;
RL Submitted (MAR-1994) to the SWISS-PROT data bank.
CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC MITOCHONDRIAL ELECTRON TRANSPORT.
CC -!- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIII FAMILY.
CC Oxidoreductase; Mitochondrion.
KW NON TER 10
FT MOD_RES 10
SQ SEQUENCE 10 AA; 1027 MW; C4E95CA33DC7633D CRC64;

Query Match 28.6%; Score 2; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 6.9e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KP 5
DB 4 KP 5

RESULT 43
GON1 CHEPR
ID GON1 CHEPR STANDARD; PRT; 10 AA.
AC P80677;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GNRH-I)
DE (Luliberin I).
OS Chelyosoma productum.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Corellidae; Chelyosoma.
ON NCBI_TaxID=71177;
RN [1]
RP SEQUENCE.
RC MEDLINE=96413669; PubMed=8816823;
RA Powell J.F.F., Reska-Skinner S.M., Prakash M.O., Fischer W.H.,
RA Park M., Rivier J.E., Craig A.G., Mackie G.O., Sherwood N.M.;
RT "Two new forms of gonadotropin-releasing hormone in a protochordate
RT and the evolutionary implications.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:10461-10464(1996).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC the secretion of both luteinizing and follicle-stimulating
CC hormones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: GNRH NEURONS LIE WITHIN BLOOD SINUSES CLOSE TO
CC THE GONADUCTS AND GONADS IN BOTH JUVENILES AND ADULTS, IMPLYING
CC THAT THE NEUROPEPTIDE IS RELEASED INTO THE BLOODSTREAM.
CC -!- MASS SPECTROMETRY: MW=1246.56; METHOD=MALDI.
CC -!- SIMILARITY: Belongs to the GNRH family.
DR InterPro; IPR002012; GNRH.
DR PROSITE; PS00473; GNRH; 1.
KW Hormone; Amidation; Pyrrolidone carboxylic acid.
FT MOD_RES 10
MOD_RES 10 10 AMIDATION (BY SIMILARITY).

SQ SEQUENCE 10 AA; 1264 MW; 284B3639DB5AB5A3 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 6.9e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KP 5
DB 8 KP 9

RESULT 44
GON1 PETMA
ID GON1 PETMA STANDARD; PRT; 10 AA.
AC P04378;
DT 20-MAR-1987 (Rel. 04, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GNRH-I)
DE (Luliberin I).
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
ON NCBI_TaxID=7757;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=86168192; PubMed=3514603;
RA Sherwood N.M., Sower S.A., Marshak D.R., Fraser B.A., Brownstein M.J.;
RT "Primary structure of gonadotropin-releasing hormone from lamprey
RT brain.";
RL J. Biol. Chem. 261:4812-4819(1986).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC the secretion of both luteinizing and follicle-stimulating
CC hormones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GNRH family.
DR PIR; A01412; RHLMSG.
DR InterPro; IPR002012; GNRH.
DR Pfam; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus; Pyrrolidone carboxylic acid.
FT MOD_RES 10
MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1244 MW; 1E4B36237B1735AB CRC64;

Query Match 28.6%; Score 2; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 6.9e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KP 5
DB 8 KP 9

RESULT 45
GON3 PETMA
ID GON3 PETMA STANDARD; PRT; 10 AA.
AC P30948;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Gonadoliberin III (Gonadotropin-releasing hormone III) (GNRH-III)
DE (Luliberin III).
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
ON NCBI_TaxID=7757;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=93178316; PubMed=8440174;
RA Sower S.A., Chiang Y.-C., Lovas S., Conlon J.M.;

```

RT "Primary structure and biological activity of a third gonadotropin-
RL releasing hormone from lamprey brain.";
RL Endocrinology 132:1125-1131(1993).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC the secretion of both luteinizing and follicle-stimulating
CC hormones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GnRH family.
DR InterPro; IPR002012; GnRH.
DR Pfam; PF00446; GnRH; 1.
DR PROSITE; PS00473; GnRH; 1.
KW Hormone; Amidation; Hypothalamus; Pyrrolidone carboxylic acid.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1277 MW; 284B36237AA1F5A3 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 6.9e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KP 5
DB 8 KP 9

RESULT 46
PNEU_HUMAN STANDARD; PRT; 10 AA.
AC P22103;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pneumadin [PNM].
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=91110910; PubMed=2274681;
RA Batra V.K., Mathur M., Mir S.A., Kapoor R., Kumar M.A.;
RT "Pneumadin: a new lung peptide which triggers antidiuresis.";
RL Regul. Pept. 30:77-87(1990).
CC -!- FUNCTION: ANTI-DIURETIC PEPTIDE THAT TRIGGERS THE RELEASE OF ADH.
CC PIR; B33143; B33143.
DR GO; GO:0030103; P:vasopressin secretion; NAS.
DR Amidation.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 956 MW; 640378DAA723276B CRC64;

Query Match 28.6%; Score 2; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 6.9e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GE 3
DB 2 GE 3

RESULT 47
PNEU_RAT STANDARD; PRT; 10 AA.
AC P21936;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pneumadin [PNM].
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]

```

```

RP SEQUENCE, AND SYNTHESIS.
RC TISSUE=Lung;
RX MEDLINE=91110910; PubMed=2274681;
RA Batra V.K., Mathur M., Mir S.A., Kapoor R., Kumar M.A.;
RT "Pneumadin: a new lung peptide which triggers antidiuresis.";
RL Regul. Pept. 30:77-87(1990).
CC -!- FUNCTION: THIS ANTI-DIURETIC PEPTIDE TRIGGERS THE RELEASE OF ADH.
CC PIR; A33143; A33143.
DR Amidation.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1048 MW; 641D00DAA723276B CRC64;

Query Match 28.6%; Score 2; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 6.9e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GE 3
DB 2 GE 3

RESULT 48
PECK_FASHE STANDARD; PRT; 10 AA.
AC P80525;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Putative phosphoenolpyruvate carboxylase [GTP] (EC 4.1.1.32)
DE (Phosphoenolpyruvate carboxylase) (PEPCK) (Newly excysted juvenile
DE protein 1) (Fragment).
OS Fasciola hepatica (Liver fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
OC Echinostomida; Echinostomata; Fasciolidae; Fasciola.
OX NCBI_TaxID=6192;
RN [1]
RP SEQUENCE.
RX MEDLINE=95366993; PubMed=7639732;
RA Tkalcovic J., Ashman K., Meeusen E.;
RT "Fasciola hepatica: rapid identification of newly excysted juvenile
RT proteins.";
RL Biochem. Biophys. Res. Commun. 213:169-174 (1995).
CC -!- CATALYTIC ACTIVITY: GTP + oxaloacetate = GDP + phosphoenolpyruvate
CC + CO(2).
CC -!- SIMILARITY: BELONGS TO THE PHOSPHOENOLPYRUVATE CARBOXYKINASE [GTP]
CC FAMILY.
DR InterPro; IPR000364; PEP carboxykin.
DR PROSITE; PS00505; PEPCK_GTP; PARTIAL.
KW Lyase; Decarboxylase; GTP-binding.
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 1069 MW; 8393A6187AA9C87A CRC64;

Query Match 28.6%; Score 2; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 6.9e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GE 3
DB 7 GE 8

RESULT 49
Q2OG_COMTE STANDARD; PRT; 10 AA.
AC P80466;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Quinolone 2-oxidoreductase, gamma chain (EC 1.3.99.17) (Fragment).
OS Comamonas testosteroni (Pseudomonas testosteroni).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Comamonadaceae; Comamonas.
OX NCBI_TaxID=285;

```

RN SEQUENCE.
 RP STRAIN=63;
 RC MEDLINE=9603589; PubMed=7556204;
 RA Schach S., Tshisuaka B., Fetner S., Lingens F.;
 RT "Quinolone 2-oxidoreductase and 2-oxo-1,2-dihydroquinoline 5,6-
 RT dioxygenase from Comamonas testosteroni 63. The first two enzymes in
 RT quinoline and 3-methylquinoline degradation.";
 RL Eur. J. Biochem. 232:536-544(1995).
 CC -1- FUNCTION: CONVERTS (3-METHYL-)QUINOLINE TO (3-METHYL-)2-OXO-
 CC 1,2-DIHYDROQUINOLINE.
 CC -1- CATALYTIC ACTIVITY: Quinolone + acceptor + H(2)O = isoquinolin-
 CC 1(2H)-one + reduced acceptor.
 CC -1- COFACTOR: FAD, MOLYBDENUM AND IRON-SULFUR.
 CC -1- PATHWAY: Degradation of quinoline and (3-methyl-)quinoline; first
 CC step.
 CC -1- SUBUNIT: HETEROHEXAMER OF TWO ALPHA CHAINS, TWO BETA CHAINS, AND
 CC TWO GAMMA CHAINS (PROBABLE).
 KW Oxidoreductase; Flavoprotein; FAD; Molybdenum.
 FT NON_TER 10 10
 FT SEQUENCE 10 AA; 1153 MW; C848CE64433B1DC6 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 6.9e+03;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EK 4
 DB 5 EK 6

RESULT 50
 RT02 BOVIN
 ID TKNB_RANCA STANDARD; PRT; 10 AA.
 AC P22689;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Mitochondrial 28S ribosomal protein S2 (MRP-S2) (Fragment).
 GN MRPS2.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 [1]
 RN SEQUENCE.
 RC TISSUE=Liver;
 RX MEDLINE=21276436; PubMed=11279123;
 RA Koc E.C., Burkhardt W., Blackburn K., Moseley A., Spremulli L.L.;
 RT "The small subunit of the mammalian mitochondrial ribosome:
 RT identification of the full complement of ribosomal proteins present.";
 RL J. Biol. Chem. 276:19363-19374(2001).
 CC -1- SUBUNIT: Component of the mitochondrial ribosome small subunit
 CC (28S) which comprises a 12S rRNA and about 30 distinct proteins.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.
 CC -1- SIMILARITY: BELONGS TO THE S2P FAMILY OF RIBOSOMAL PROTEINS.
 CC InterPro: IPR001865; Ribosomal S2.
 DR Pfam: PF00318; Ribosomal S2; PARTIAL.
 DR PROSITE: PS00962; RIBOSOMAL_S2_1; PARTIAL.
 KW Ribosomal protein; Mitochondrion.
 FT NON_TER 10 10
 FT SEQUENCE 10 AA; 1246 MW; 6A7A679C04B476B CRC64;

Query Match 28.6%; Score 2; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 6.9e+03;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PY 6
 DB 4 PY 5

RESULT 51
 SPI HALRO
 ID SPI_HALRO STANDARD; PRT; 10 AA.
 AC Q10397;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Serine proteinase inhibitor (Fragment).
 OS Halocynthia roretzi (Sea squirt).
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
 OC Stolidobranchia; Pyridae; Halocynthia.
 OC NCBI_TaxID=7729;
 [1]
 RN SEQUENCE.
 RC TISSUE=Hemolymph;
 RX MEDLINE=96321313; PubMed=8759295;
 RA Shishikura F., Abe T., Ohtake S.-I., Tanaka K.;
 RT "Purification and characterization of a 58,000-Da proteinase
 RT inhibitor from the hemolymph of a solitary ascidian, Halocynthia
 RT roretzi.";
 RL Comp. Biochem. Physiol. 114B:1-9(1996).
 CC -1- FUNCTION: STRONGLY INHIBITS TRYPSIN AND PLASMA ENZYME(S) ACTIVITY.
 CC -1- SUBUNIT: Monomer.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 DR InterPro: IPR000215; Serpin.
 DR PROSITE: PS00284; SERPIN; PARTIAL.
 KW Serpin; Serine protease inhibitor; Glycoprotein; Plasma.
 FT NON_TER 10 10
 FT SEQUENCE 10 AA; 1104 MW; 4225C73B1B187AA3 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 6.9e+03;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GE 3
 DB 5 GE 6

RESULT 52
 TKNB_RANCA
 ID TKNB_RANCA STANDARD; PRT; 10 AA.
 AC P22689;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ranatachikin B (RTK B).
 OS Rana catesbeiana (Bull frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranidae; Rana.
 OC NCBI_TaxID=8400;
 [1]
 RN SEQUENCE, AND SYNTHESIS.
 RC TISSUE=Brain;
 RX MEDLINE=91254337; PubMed=2043143;
 RA Kozawa H., Hino J., Minamino N., Kangawa K., Matsuo H.;
 RT "Isolation of four novel tachykinins from frog (Rana catesbeiana)
 RT brain and intestine.";
 RL Biochem. Biophys. Res. Commun. 177:588-595(1991).
 [2]
 RN SEQUENCE.
 RC TISSUE=Intestine;
 RX MEDLINE=9403216; PubMed=8210506;
 RA Kangawa K., Kozawa H., Hino J., Minamino N., Matsuo H.;
 RT "Four novel tachykinins in frog (Rana catesbeiana) brain and
 RT intestine.";
 RL Regul. Pept. 46:81-88(1993).
 CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
 CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
 CC MUSCLES.
 CC -1- SUBCELLULAR LOCATION: Secreted.

CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
 DR PIR; B61033; B61033.
 DR InterPro; IPR002040; Tachykinin.
 DR PROSITE; PS00267; TACHYKININ; 1.
 KW Tachykinin; Neuropeptide; Amidation.
 FT MOD_RES 10 10
 SQ SEQUENCE 10 AA; 1210 MW; 917556B59D5BAB5 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 6.9e+03;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 YK 7
 ||
 Db 1 YK 2

RESULT 53

ID_TKS1_AEDAE STANDARD; PRT; 10 AA.
 AC P42634;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Sialokinin I.
 OS Aedes aegypti (Yellowfever mosquito).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes.
 OX NCBI_TaxID=7159;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=Rockefeller; TISSUE=Salivary gland;
 RA MEDLINE=94105119; PubMed=8278354;
 RA Champagne D.E., Ribeiro J.M.C.;
 RT "Sialokinin I and II: vasodilatory tachykinins from the yellow fever
 RT mosquito Aedes aegypti.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:138-142(1994).
 CC -!- FUNCTION: VASODILATORY PEPTIDE. MAY ACTIVATE MACROPHAGES AT THE
 CC SITE OF FEEDING.

CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
 DR PIR; A49581; A49581.
 DR InterPro; IPR002040; Tachykinin.
 DR PROSITE; PS00267; TACHYKININ; 1.
 KW Tachykinin; Neuropeptide; Amidation.
 FT MOD_RES 10 10
 SQ SEQUENCE 10 AA; 1145 MW; 3DCFDEB59C33AA8 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 6.9e+03;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TG 2
 ||
 Db 2 TG 3

RESULT 54

ID_TKS2_AEDAE STANDARD; PRT; 10 AA.
 AC P42635;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Sialokinin II.
 OS Aedes aegypti (Yellowfever mosquito).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes.
 OX NCBI_TaxID=7159;
 RN [1]
 RP SEQUENCE.

RC STRAIN=Rockefeller; TISSUE=Salivary gland;
 RX MEDLINE=94105119; PubMed=8278354;

RA Champagne D.E., Ribeiro J.M.C.;
 RT "Sialokinin I and II: vasodilatory tachykinins from the yellow fever
 RT mosquito Aedes aegypti.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:138-142(1994).
 CC -!- FUNCTION: VASODILATORY PEPTIDE. MAY ACTIVATE MACROPHAGES AT THE
 CC SITE OF FEEDING.

CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
 DR PIR; B49581; B49581.
 DR InterPro; IPR002040; Tachykinin.
 DR PROSITE; PS00267; TACHYKININ; 1.
 KW Tachykinin; Neuropeptide; Amidation.
 FT MOD_RES 10 10
 SQ SEQUENCE 10 AA; 1146 MW; 3DC77C6B59C33AA8 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 6.9e+03;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TG 2
 ||
 Db 2 TG 3

RESULT 55

ID_URAL_HUMAN STANDARD; PRT; 10 AA.
 AC P32118;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Unknown protein from 2D-page of red blood cells (Spot 1) (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.

RC TISSUE=Erythrocyte;
 RX MEDLINE=94147970; PubMed=8313871;
 RA Golaz O., Hughes G.J., Frutiger S., Paquet N., Bairoch A.,
 RA Pasquali C., Sanchez J.-C., Tissot J.-D., Appel R.D., Walzer C.,
 RA Balant L., Hochstrasser D.P.;
 RT "Plasma and red blood cell protein maps: update 1993.";
 RL Electrophoresis 14:1223-1231(1993).
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 5.4, ITS MW IS: 23 kDa.
 DR SWISS-2DPAGE; P32118; HUMAN.
 FT NON_TER 1 1
 FT NON_TER 10 10
 SQ SEQUENCE 10 AA; 977 MW; 723C6581ADD0587B CRC64;

Query Match 28.6%; Score 2; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 6.9e+03;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TG 2
 ||
 Db 8 TG 9

RESULT 56

ID_UR7_HUMAN STANDARD; PRT; 10 AA.
 AC P34990;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Unknown protein from 2D-page of red blood cells (Spot 2D-006HO)
 DE (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;
 [1]
 RP SEQUENCE.
 RC TISSUE=Erythrocyte;
 RA Golaz O., Hughes G.J., Frutiger S., Paquet N., Bairoch A.,
 RA Pasquali C., Sanchez J.-C., Tissot J.-D., Appel R.D., Walzer C.,
 RA Balant L., Hochstrasser D.F.;
 RL Submitted (FEB-1994) to the SWISS-PROT data bank.
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 4.77, ITS MW IS: 26 kDa.
 DR SWISS-2DPAGE; P34990; HUMAN.
 FT NON TER 10
 SQ SEQUENCE 10 AA; 1161 MW; F72D82E9C44B1871 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 6.9e+03;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GE 3
 ||
 Db 4 GE 5

RESULT 57
 ID_URE3 MORMO STANDARD; PRT; 10 AA.
 AC P17339;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DE Urease gamma subunit (EC 3.5.1.5) (6 kDa subunit) (Urea
 DE amidohydrolase) (Fragment).
 GN UREA.
 OS Morganella morganii (Proteus morganii).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Morganella.
 OX NCBI_TaxID=582;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=90264298; PubMed=2345135;
 RA Hu L.-T., Nicholson E.B., Jones B.D., Lynch M.J., Mobley H.L.T.;
 RT "Morganella morganii urease: purification, characterization, and
 RT isolation of gene sequences."
 RL J. Bacteriol. 172:3073-3080(1990).
 CC -!- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).
 CC -!- SUBUNIT: (ALPHA, BETA, GAMMA) (3) (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE UREASE GAMMA SUBUNIT FAMILY.
 DR PIR; C35389; C35389.
 KW Hydrolase.
 FT NON TER 10
 SQ SEQUENCE 10 AA; 1171 MW; 4B313BCB07771A7 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 6.9e+03;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EK 4
 ||
 Db 9 EK 10

RESULT 58
 ID_VEG6 BACSU STANDARD; PRT; 10 AA.
 AC P80639;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Vegetative protein 6 (VEG6) (Fragment).
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]

RP SEQUENCE.
 RC STRAIN=168 / IS58;
 RX MEDLINE=97237728; PubMed=9084183;
 RA Schmid R., Bernhardt J., Antelmann H., Voelker U., Mach H.,
 RA Voelker A., Hecker M.;
 RT "Identification of vegetative proteins for a two-dimensional protein
 RT index of Bacillus subtilis."
 RL Microbiology 143:991-998(1997).
 CC -!- CAUTION: Could not be found in the genome of B. subtilis 168.
 FT NON TER 10
 SQ SEQUENCE 10 AA; 973 MW; 8793A6B2C8772861 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 6.9e+03;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TG 2
 ||
 Db 3 TG 4

RESULT 59
 ID_GRWV HUMAN STANDARD; PRT; 3 AA.
 AC P01157;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 21-JUL-1986 (Rel. 01, Last annotation update)
 DE Growth-modulating peptide.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=77162369; PubMed=858356;
 RA Schleisinger D.H., Pickart L., Thaler M.M.;
 RT "Growth-modulating serum tripeptide is glycyl-histidyl-lysine."
 RL Experientia 33:324-325(1977).
 CC -!- MISCELLANEOUS: THIS SERUM TRIPEPTIDE HAS BEEN FOUND TO STIMULATE
 CC GROWTH OF SOME CELL TYPES AND TO INHIBIT OTHER TYPES IN VITRO.
 DR GO; GO:0001558; P:regulation of cell growth; NAS.
 SQ SEQUENCE 3 AA; 340 MW; 6331B81000000000 CRC64;

Query Match 14.3%; Score 1; DB 1; Length 3;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 G 2
 |
 Db 1 G 1

RESULT 60
 ID_LUXE VIBFI STANDARD; PRT; 3 AA.
 AC P24272;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Long-chain-fatty-acid--luciferin-component ligase (EC 6.2.1.19) (Acyl-
 DE protein synthetase) (Fragment).
 GN LUXE.
 OS Vibrio fischeri.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxID=668;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91072226; PubMed=2254256;
 RA Swartman E., Kapoor S., Graham A.F., Meighen E.A.;
 RT "A new Vibrio fischeri lux gene precedes a bidirectional termination
 RT site for the lux operon.";

RL J. Bacteriol. 172:6797-6802(1990).
 CC -!- FUNCTION: ACYL-PROTEIN SYNTHETASE ACTIVATES TETRADECANOIC ACID.
 CC IT IS A COMPONENT OF THE FATTY ACID REDUCTASE COMPLEX RESPONSIBLE
 CC FOR CONVERTING TETRADECANOIC ACID TO THE ALDEHYDE WHICH SERVES AS
 CC SUBSTRATE IN THE LUCIFERASE-CATALYZED REACTION.
 CC -!- CATALYTIC ACTIVITY: ATP + an acid + protein = AMP + diphosphate +
 CC an acyl-protein thioester.
 CC -!- PATHWAY: Bioluminescent fatty acid reduction system; second step.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M62812; -, NOT_ANNOTATED_CDS.
 CC Luminescence; Ligase.
 KW NON TER 1
 FT SEQUENCE 3 AA; 374 MW; 6AA3303000000000 CRC64;
 SQ
 Query Match 14.3%; Score 1; DB 1; Length 3;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 4 K 4
 Db 2 K 2
 RESULT 61
 ID THYL_PIG STANDARD; PRT; 3 AA.
 AC P01151;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Thyroliberin (Thyrotropin releasing hormone) (TRH) (Protirelin).
 OS Sus scrofa (Pig),
 OS Ovis aries (Sheep),
 OS Bombina orientalis (Oriental fire-bellied toad), and
 OS Notophthalmus viridescens (Eastern newt) (Triturus viridescens).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823, 9940, 8346, 8316;
 RN [1]
 RP SEQUENCE.
 RC SPECIES=Pig; TISSUE=Hypothalamus;
 RA MEDLINE=70136150; PubMed=4984938;
 RX Nair R.M.G., Barrett J.F., Bowers C.Y., Schally A.V.;
 RT "Structure of porcine thyrotropin releasing hormone.";
 RL Biochemistry 9:1103-1106(1970).
 RN [2]
 RP SEQUENCE.
 RC SPECIES=Pig;
 RA MEDLINE=70039904; PubMed=4982117;
 RX Boller J., Enzmann F., Folkers K., Bowers C.Y., Schally A.V.;
 RT "The identity of chemical and hormonal properties of the thyrotropin
 RT releasing hormone and pyroglutamyl-histidyl-proline amide.";
 RL Biochem. Biophys. Res. Commun. 37:705-710(1969).
 RN [3]
 RP SEQUENCE.
 RC SPECIES=Sheep; TISSUE=Hypothalamus;
 RA Desiderio D.M. Jr., Burgess R., Dunn T.F., Vale W., Guillemin R.,
 RA Ward D.N.;
 RT "The elucidation of the primary structure of the hypothalamic thyroid
 RT stimulating hormone releasing factor of ovine origin by means of mass
 RT spectrometry.";
 RL Org. Mass Spectrom. 5:221-228(1971).
 RN [4]
 RP SYNTHESIS.
 RC SPECIES=Sheep;

RX MEDLINE=70163386; PubMed=4985794;
 RA Burgess R., Dunn T.F., Desiderio D.M., Ward D.N., Vale W.,
 RA Guillemin R.;
 RT "Characterization of ovine hypothalamic hypophysiotropic
 RT TSH-releasing factor.";
 RL Nature 226:321-325(1970).
 RN [5]
 RP SEQUENCE.
 RC SPECIES=B.orientalis; TISSUE=Skin;
 RX MEDLINE=76138399; PubMed=815011;
 RA Yasuhara T., Nakajima T.;
 RT "Letter: Occurrence of Pyr-His-Pro-NH2 in the frog skin.";
 RL Chem. Pharm. Bull. 23:3301-3303(1975).
 RN [6]
 RP SEQUENCE.
 RC SPECIES=N.viridescens;
 RX MEDLINE=75035605; PubMed=4214528;
 RA Grimm-Joergensen Y., McKelvy J.F.;
 RT "Biosynthesis of thyrotropin releasing factor by newt (Triturus
 RT viridescens) brain in vitro. Isolation and characterization of
 RT thyrotropin releasing factor.";
 RL J. Neurochem. 23:471-478(1974).
 CC -!- FUNCTION: TRH FUNCTIONS AS A REGULATOR OF THE BIOSYNTHESIS OF TSH
 CC IN THE ANTERIOR PITUITARY GLAND AND AS A NEUROTRANSMITTER/
 CC NEUROMODULATOR IN THE CENTRAL AND PERIPHERAL NERVOUS SYSTEMS.
 CC
 DR PIR; A90919; RHTDFO.
 DR PIR; A92971; A92971.
 DR PIR; A91750; RSHST.
 KW Amidation; Pyrrolidone carboxylic acid.
 FT MOD RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD RES 3 3 AMIDATION.
 SQ SEQUENCE 3 AA; 380 MW; 7761F6B000000000 CRC64;
 Query Match 14.3%; Score 1; DB 1; Length 3;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 5 P 5
 Db 3 P 3
 RESULT 62
 ID ACH1_ACHFU STANDARD; PRT; 4 AA.
 AC P35904;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Achatin-I.
 OS Achatina fulica (Giant African snail).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
 OC Sigmurethra; Achatinoidea; Achatinidae; Achatina.
 OX NCBI_TaxID=6530;
 RN [1]
 RP SEQUENCE, CHARACTERIZATION, AND SYNTHESIS.
 RC STRAIN=Perussac; TISSUE=Ganglion;
 RX MEDLINE=89273551; PubMed=2597281;
 RA Kanatani Y., Minakata H., Kenny P.T.M., Iwashita T., Watanabe K.,
 RA Nunase K., Sun X.P., Yongsiri A., Kim K.H., Novales-Li P.,
 RA Novales E.T., Kanapi C.G., Takeuchi H., Nomoto K.;
 RT "Achatin-I, an endogenous neuroexcitatory tetrapeptide from Achatina
 RT fulica Perussac containing a D-amino acid residue.";
 RL Biochem. Biophys. Res. Commun. 160:1015-1020(1989).
 RN [2]
 RP CHARACTERIZATION.
 RC STRAIN=Perussac; TISSUE=Heart atrium;
 RX MEDLINE=91264856; PubMed=1675568;
 RA Fujimoto K., Kubota I., Yasuda-Kamatani Y., Minakata H., Nomoto K.,
 RA Yoshida M., Harada A., Muneoka Y., Kobayashi M.;
 RT "Purification of achatin-I from the atria of the African giant snail,
 RT Achatina fulica, and its possible function.";
 RL Biochem. Biophys. Res. Commun. 177:847-853(1991).


```

RN X-RAY CRYSTALLOGRAPHY.
RP MEDLINE=9301014529; PubMed=1399265;
RA Ishida T., In Y., Doi M., Inoue M., Yasuda-Kamatani Y., Minakata H.,
RA Iwashita T., Nomoto K.;
RT "Crystal structure and molecular conformation of achain-I
RT (H-Gly-D-Phe-Ala-Asp-OH), an endogenous neuropeptide containing a
RT D-amino acid residue.";
RL Int. J. Pept. Protein Res. 39:258-264(1992).
CC -!- FUNCTION: NEUROEXCITATORY PEPTIDE; INCREASES THE IMPULSE FREQUENCY
CC AND PRODUCES A SPIKE BROADENING OF THE IDENTIFIED HEART EXCITATORY
CC NEURON (PON); ALSO ENHANCES THE AMPLITUDE AND FREQUENCY OF THE
CC HEART BEAT. HAS ALSO AN EFFECT ON SEVERAL OTHER MUSCLES.
DR PIR; A32480; A32480.
KW Hormone; D-amino acid.
FT MOD_RES 2 2 D-PHENYLALANINE.
SQ SEQUENCE 4 AA; 408 MW; 6AADD9C8100000000 CRC64;

Query Match 14.3%; Score 1; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 G 2
EB 1 G 1

RESULT 63
DCML_PSECH STANDARD; PRT; 4 AA.
AC P19916;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Carbon monoxide dehydrogenase large chain (EC 1.2.99.2) (CO
DE dehydrogenase subunit L) (CO-DH L) (Fragment).
GN CUTL.
OS Pseudomonas carboxydohydrogena.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae.
OX NCBI_TaxID=290;
RN [1]
RP SEQUENCE.
RX MEDLINE=90055678; PubMed=2818128;
RA Kraut M., Hugendieck I., Herwig S., Meyer O.;
RT "Homology and distribution of CO dehydrogenase structural genes in
RT carboxydotropic bacteria.";
RL Arch. Microbiol. 152:335-341(1989).
CC -!- FUNCTION: Catalyzes the oxidation of carbon monoxide to carbon
CC dioxide.
CC -!- CATALYTIC ACTIVITY: CO + H(2)O + acceptor = CO(2) + reduced
CC acceptor.
CC -!- COFACTOR: BINDS TWO 2FE-2S CLUSTERS.
CC -!- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
CC SMALL.
DR PIR; P19146; P19146.
KW Oxidoreductase; Metal-binding; Iron-sulfur; Iron; 2Fe-2S.
FT NON_TER 4 4
SQ SEQUENCE 4 AA; 420 MW; 6DD33DD6F00000000 CRC64;

Query Match 14.3%; Score 1; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 K 4
DB 3 K 3

RESULT 65
EOSI_HUMAN STANDARD; PRT; 4 AA.
AC P02731;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1986 (Rel. 01, Last annotation update)
DE Eosinophilotoxic peptides.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=76078412; PubMed=1060093;
RA Goetzl E.J., Austen K.F.;
RT "Purification and synthesis of eosinophilotoxic tetrapeptides of
RT human lung tissue: identification as eosinophil chemotactic factor of
RT anaphylaxis.";
RL Proc. Natl. Acad. Sci. U.S.A. 72:4123-4127(1975).
CC -!- MISCELLANEOUS: THESE PEPTIDES ARE RELEASED FROM MAST CELLS IN LUNG
CC (AND OTHER TISSUES) DURING HYPERSENSITIVITY REACTIONS
CC (ANAPHYLAXIS). THEIR ACTIVITIES, PREFERENTIALLY AFFECTING
CC EOSINOPHILS, INCLUDE CHEMOTAXIS, CHEMOTACTIC DEACTIVATION, RELEASE
CC OF ENZYMES, AND STIMULATION OF THE HEXOSE MONOPHOSPHATE SHUNT.
DR GO; GO:0030105; P:anaphylaxis; IDA.
DR GO; GO:0006935; P:chemotaxis; IDA.
FT VARIANT 1 1 V->A (IN OTHER PEPTIDE).
FT /FTID=VAR_005201.
SQ SEQUENCE 4 AA; 390 MW; 6B05B862A00000000 CRC64;

Query Match 14.3%; Score 1; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 G 2
DB 2 G 2

RESULT 64
DCMS_PSECH STANDARD; PRT; 4 AA.
ID P19918;

```

```

DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Carbon monoxide dehydrogenase small chain (EC 1.2.99.2) (CO
DE dehydrogenase subunit S) (CO-DH S) (Fragment).
GN CUTS.
OS Pseudomonas carboxydohydrogena.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae.
OX NCBI_TaxID=290;
RN [1]
RP SEQUENCE.
RX MEDLINE=90055678; PubMed=2818128;
RA Kraut M., Hugendieck I., Herwig S., Meyer O.;
RT "Homology and distribution of CO dehydrogenase structural genes in
RT carboxydotropic bacteria.";
RL Arch. Microbiol. 152:335-341(1989).
CC -!- FUNCTION: Catalyzes the oxidation of carbon monoxide to carbon
CC dioxide.
CC -!- CATALYTIC ACTIVITY: CO + H(2)O + acceptor = CO(2) + reduced
CC acceptor.
CC -!- COFACTOR: BINDS TWO 2FE-2S CLUSTERS.
CC -!- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
CC SMALL.
DR PIR; P19146; P19146.
KW Oxidoreductase; Metal-binding; Iron-sulfur; Iron; 2Fe-2S.
FT NON_TER 4 4
SQ SEQUENCE 4 AA; 420 MW; 6DD33DD6F00000000 CRC64;

Query Match 14.3%; Score 1; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 K 4
DB 3 K 3

RESULT 65
EOSI_HUMAN STANDARD; PRT; 4 AA.
AC P02731;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1986 (Rel. 01, Last annotation update)
DE Eosinophilotoxic peptides.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=76078412; PubMed=1060093;
RA Goetzl E.J., Austen K.F.;
RT "Purification and synthesis of eosinophilotoxic tetrapeptides of
RT human lung tissue: identification as eosinophil chemotactic factor of
RT anaphylaxis.";
RL Proc. Natl. Acad. Sci. U.S.A. 72:4123-4127(1975).
CC -!- MISCELLANEOUS: THESE PEPTIDES ARE RELEASED FROM MAST CELLS IN LUNG
CC (AND OTHER TISSUES) DURING HYPERSENSITIVITY REACTIONS
CC (ANAPHYLAXIS). THEIR ACTIVITIES, PREFERENTIALLY AFFECTING
CC EOSINOPHILS, INCLUDE CHEMOTAXIS, CHEMOTACTIC DEACTIVATION, RELEASE
CC OF ENZYMES, AND STIMULATION OF THE HEXOSE MONOPHOSPHATE SHUNT.
DR GO; GO:0030105; P:anaphylaxis; IDA.
DR GO; GO:0006935; P:chemotaxis; IDA.
FT VARIANT 1 1 V->A (IN OTHER PEPTIDE).
FT /FTID=VAR_005201.
SQ SEQUENCE 4 AA; 390 MW; 6B05B862A00000000 CRC64;

Query Match 14.3%; Score 1; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 G 2
DB 2 G 2

RESULT 64
DCMS_PSECH STANDARD; PRT; 4 AA.
ID P19918;

```

QY 2 G 2
Db 2 G 2

RESULT 66

FAF3_HIRME ID_FAF3_HIRME STANDARD; PRT; 4 AA.
AC P42562;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE FMRamide-like neuropeptide YLRF-amide.
OS Hirudo medicinalis (Medicinal leech).
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC Arynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
NCBI_TaxID=6421;
RN [1]
RP SEQUENCE.

RA MEDLINE=92195954; PubMed=1686933;
RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT "Identification of Rfamide neuropeptides in the medicinal leech."
RL Peptides 12:897-908(1991).
CC -!- FAMILY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.

KW Neuropeptide; Amidation.
FT MOD_RES 4 4 AMIDATION.
SQ SEQUENCE 4 AA; 598 MW; 69D4073B30000000 CRC64;

Query Match 14.3%; Score 1; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 Y 6
Db 1 Y 1

RESULT 67

FAR4_HIRME ID_FAR4_HIRME STANDARD; PRT; 4 AA.
AC P42563;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE FMRamide-like neuropeptide YMRP-amide.
OS Hirudo medicinalis (Medicinal leech).
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC Arynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
NCBI_TaxID=6421;
RN [1]
RP SEQUENCE.

RA MEDLINE=92195954; PubMed=1686933;
RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT "Identification of Rfamide neuropeptides in the medicinal leech."
RL Peptides 12:897-908(1991).
CC -!- FAMILY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.

KW Neuropeptide; Amidation.
FT MOD_RES 4 4 AMIDATION.
SQ SEQUENCE 4 AA; 616 MW; 69D4068B30000000 CRC64;

Query Match 14.3%; Score 1; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 Y 6
Db 1 Y 1

RESULT 68

FFKA_ANTEL

ID FFKA_ANTEL STANDARD; PRT; 4 AA.
AC P58706;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Antho-Kkamide.
DE Antho-Kkamide.
OS Anthopleura elegantissima (Sea anemone).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
OC Nynanthae; Actiniidae; Anthopleura.
NCBI_TaxID=6110;
RN [1]
RP SEQUENCE.

RA MEDLINE=92028852; PubMed=1681803;
RA Nothacker H.-P., Rinehart K.L. Jr., Grimmelikhuijzen C.J.P.;
RT "Isolation of L-3-phenyllactyl-Phe-Lys-Ala-NH2 (Antho-Kkamide), a novel neuropeptide from sea anemones."
RL Biochem. Biophys. Res. Commun. 179:1205-1211(1991).
RN [2]
RP FUNCTION.

RA MEDLINE=93391436; PubMed=8397415;
RA McFarlane I.D., Hudman D., Nothacker H.-P., Grimmelikhuijzen C.J.P.;
RT "The expansion behaviour of sea anemones may be coordinated by two inhibitory neuropeptides, Antho-Kkamide and Antho-Riamide."
RL Proc. R. Soc. Lond., B, Biol. Sci. 253:183-188(1993).

CC -!- FUNCTION: Inhibits spontaneous contractions in several muscle groups. May be involved in the expansion phase of feeding behaviour in sea anemones.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: Neuron-specific.

DR PIR; JQ1273; JQ1273.

KW Neuropeptide; Amidation.

FT MOD_RES 1 1 L-3-PHENYLLACTYL.

FT MOD_RES 4 4 AMIDATION.

SQ SEQUENCE 4 AA; 512 MW; 6DD339C9A0000000 CRC64;

Query Match 14.3%; Score 1; DB 1; Length 4;

Best Local Similarity 100.0%; Pred. No. 1.3e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 K 4
Db 3 K 3

RESULT 69

FYRI_ANTEL ID_FYRI_ANTEL STANDARD; PRT; 4 AA.
AC P58706;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Antho-Riamide I [Contains: Antho-Riamide II].
OS Anthopleura elegantissima (Sea anemone).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
OC Nynanthae; Actiniidae; Anthopleura.
NCBI_TaxID=6110;
RN [1]
RP SEQUENCE.

RA MEDLINE=92270459; PubMed=1821096;

RA Nothacker H.-P., Rinehart K.L. Jr., McFarlane I.D.,

RA Grimmelikhuijzen C.J.P.;

RT "Isolation of two novel neuropeptides from sea anemones: the unusual, biologically active L-3-phenyllactyl-Tyr-Arg-Ile-NH2 and its des-phenyllactyl fragment Tyr-Arg-Ile-NH2."

RL Peptides 12:1165-1173(1991).

RN [2]
RP FUNCTION.

RA MEDLINE=93391436; PubMed=8397415;

RA McFarlane I.D., Hudman D., Nothacker H.-P., Grimmelikhuijzen C.J.P.;
RT "The expansion behaviour of sea anemones may be coordinated by two inhibitory neuropeptides, Antho-Kkamide and Antho-Riamide."

RL Proc. R. Soc. Lond., B, Biol. Sci. 253:183-188(1993).

CC -!- FUNCTION: Inhibits spontaneous contractions in several muscle groups.

CC groups. May be involved in the expansion phase of feeding
 CC behaviour in sea anemones.
 CC -|- SUBCELLULAR LOCATION: Secreted.
 CC -|- TISSUE SPECIFICITY: Neuron-specific.
 KW Neuropeptide; Amidation.
 FT CHAIN 1 4 ANTHO-RIAMIDE I.
 FT CHAIN 2 4 ANTHO-RIAMIDE II.
 FT MOD RES 1 1 L-3-PHENYLACTYL.
 FT MOD RES 4 4 AMIDATION.
 SQ SEQUENCE 4 AA; 598 MW; 60441B59A0000000 CRC64;

Query Match 14.3%; Score 1; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 Y 6
 DB 2 Y 2

RESULT 70
 OCPI_OCTMI
 ID OCPI_OCTMI STANDARD; PRT; 4 AA.
 AC P58618;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cardioactive peptides Ocp-1/Ocp-2.
 OS Octopus minor (Octopus).
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;
 OC Octopodiformes; Octopoda; Incirrata; Octopodidae; Octopus.
 OX NCBI_TaxID=89766;
 RN [1]
 RP SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.
 RC TISSUE=Brain;
 RX MEDLINE=20336815; PubMed=10876044;
 RA Iwakoshi E., Hisada M., Minakata H.;
 RT "Cardioactive peptides isolated from the brain of a Japanese octopus,
 RL Octopus minor";
 RL Peptides 21:623-630(2000).
 CC -|- FUNCTION: Cardioactive; has both positive chronotropic and
 CC inotropic effects on the heart. Ocp-2 is a 1000 time less
 CC active than Ocp-1.
 CC -|- SUBCELLULAR LOCATION: Secreted.
 CC -|- PTM: Ocp-2 has L-Phe instead of D-Phe.
 CC -|- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI.
 KW Hormone; D-amino acid.
 FT MOD RES 2 2 D-PHENYLALANINE.
 SQ SEQUENCE 4 AA; 394 MW; 6AA879C810000000 CRC64;

Query Match 14.3%; Score 1; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 G 2
 DB 1 G 1

RESULT 71
 OCPI_OCTMI
 ID OCPI_OCTMI STANDARD; PRT; 4 AA.
 AC P58619;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cardioactive peptides Ocp-3/Ocp-4.
 OS Octopus minor (Octopus).
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;
 OC Octopodiformes; Octopoda; Incirrata; Octopodidae; Octopus.
 OX NCBI_TaxID=89766;
 RN [1]
 RP SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.

RC TISSUE=Brain;
 RX MEDLINE=20336815; PubMed=10876044;
 RA Iwakoshi E., Hisada M., Minakata H.;
 RT "Cardioactive peptides isolated from the brain of a Japanese octopus,
 RL Octopus minor";
 RL Peptides 21:623-630(2000).
 CC -|- FUNCTION: Cardioactive; has both positive chronotropic and
 CC inotropic effects on the heart. Ocp-4 is a 1000 time less
 CC active than Ocp-3.
 CC -|- SUBCELLULAR LOCATION: Secreted.
 CC -|- PTM: Ocp-4 has D-Ser instead of L-Ser.
 CC -|- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI.
 KW Hormone; D-amino acid.
 FT MOD RES 2 2 D-SERINE (IN OCP-4).
 SQ SEQUENCE 4 AA; 463 MW; 6AB365B810000000 CRC64;

Query Match 14.3%; Score 1; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 G 2
 DB 1 G 1

RESULT 72
 RM01_YEAST
 ID RM01_YEAST STANDARD; PRT; 4 AA.
 AC P36515;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE Mitochondrial 60S ribosomal protein L1 (YmL1) (Fragment).
 GN MRPL1.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=91285106; PubMed=2060626;
 RA Grohmann L., Graack H.-R., Kruff V., Choli T., Goldschmidt-Reisin S.,
 RA Kitakawa M.;
 RT "Extended N-terminal sequencing of proteins of the large ribosomal
 RL subunit from yeast mitochondria";
 RL FEBS Lett. 284:51-56(1991).
 DR PIR; S17255; S17255.
 DR SGD; L0002681; MRPL1.
 KW Ribosomal protein; Mitochondrion.
 FT NON TER 4 4
 SQ SEQUENCE 4 AA; 402 MW; 7771B2D5D0000000 CRC64;

Query Match 14.3%; Score 1; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 T 1
 DB 3 T 3

RESULT 73
 AL14_CARMA
 ID AL14_CARMA STANDARD; PRT; 5 AA.
 AC P81817;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Carcinustatin 14.
 OS Carcinus maenas (Common shore crab) (Green crab).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
 OC Eubrachyura; Portunioidea; Portunidae; Carcinus.

OX NCBI_TaxID=6759;
 RN [1]
 RP SEQUENCE
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
 RX MEDLINE=98121193; PubMed=9461295;
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
 RA Thorpe A.;
 RT "Isolation and identification of multiple neuropeptides of the
 RT allatostatin superfamily in the shore crab *Carcinus maenas*.";
 RL Eur. J. Biochem. 250:727-734(1997).
 CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 KW Neuropeptide; Amidation; Multigene family.
 FT MOD RES 5 5 AMIDATION (POTENTIAL).
 SQ SEQUENCE 5 AA; 586 MW; 672879D5AB300000 CRC64;

 Query Match 14.3%; Score 1; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 OY 2 G 2
 DB 4 G 4

 RESULT 74
 B10A CITR STANDARD; PRT; 5 AA.
 ID B10A CITR STANDARD; PRT; 5 AA.
 AC P13071;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Adenosylmethionine-8-amino-7-oxononanoate aminotransferase
 DE (EC 2.6.1.62) (7,8-diamino-pelargonic acid aminotransferase) (DAPA
 DE aminotransferase) (Fragment).
 GN B10A.
 OS Citrobacter freundii.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Citrobacter.
 OX NCBI_TaxID=546;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89006280; PubMed=2971595;
 RA Shuan D., Campbell A.;
 RT "Transcriptional regulation and gene arrangement of *Escherichia coli*,
 RT Citrobacter freundii and *Salmonella typhimurium* biotin operons.";
 RL Gene 67:203-211(1988).
 CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 8-amino-7-oxononanoate = S-adenosyl-4-methylthio-2-oxobutanoate + 7,8-diaminononanoate.
 CC -!- COFACTOR: Pyridoxal phosphate.
 CC -!- PATHWAY: Biotin biosynthesis.
 CC -!- SUBUNIT: Homodimer.
 CC -!- SIMILARITY: Belongs to class-III of pyridoxal-phosphate-dependent aminotransferases.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL; M21922; -; NOT_ANNOTATED_CDS.
 DR PIR; I40697; I40697.
 DR InterPro; IPR005814; Aminotrans_3.
 DR PROSITE; PS00600; AA_TRANSFER_CLASS_3; PARTIAL.
 KW Biotin biosynthesis; Transferase; Aminotransferase;
 KW Pyridoxal phosphate.
 FT NON TER 5 5
 SQ SEQUENCE 5 AA; 582 MW; 6AAAB1B1A6F00000 CRC64;

Query Match 14.3%; Score 1; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 OY 1 T 1
 DB 2 T 2

 RESULT 75
 BPP7 BOTIN STANDARD; PRT; 5 AA.
 ID BPP7 BOTIN STANDARD; PRT; 5 AA.
 AC P30425;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Bradykinin-potentiating peptide S5,2 (5A) (Angiotensin-converting enzyme inhibitor).
 DE Bothrops insularis (Island jararaca) (Queimada jararaca).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
 OC Viperidae; Crotalinae; Bothrops.
 OX NCBI_TaxID=8723;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Venom;
 RX MEDLINE=90351557; PubMed=2386615;
 RA Cintra A.C.O., Vieira C.A., Giglio J.R.;
 RT "Primary structure and biological activity of bradykinin potentiating peptides from *Bothrops insularis* snake venom.";
 RL J. Protein Chem. 9:221-227(1990).
 CC -!- FUNCTION: This peptide both inhibits the activity of the angiotensin-converting enzyme and enhances the action of bradykinin by inhibiting the kinases that inactivate it.
 CC It acts as an indirect hypotensive agent.
 DR PIR; G37196; G37196.
 KW Hypotensive agent; Pyrrolidone carboxylic acid.
 FT MOD RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 SQ SEQUENCE 5 AA; 629 MW; 776DC37326B00000 CRC64;

 Query Match 14.3%; Score 1; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 OY 4 K 4
 DB 2 K 2

 RESULT 76
 EI03 LITRU STANDARD; PRT; 5 AA.
 ID EI03 LITRU STANDARD; PRT; 5 AA.
 AC P82039;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Electrin 3.
 OS Litoria rubella (Desert tree frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
 OC Pelodyadinae; Litoria.
 OX NCBI_TaxID=104895;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Skin secretion;
 RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
 RT "Peptides from the skin glands of the Australian buzzing tree frog *Litoria electrica*. Comparison with the skin peptides from *Litoria rubella*.";
 RL Aust. J. Chem. 52:639-645(1999).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Skin.
 KW Amphibian defense peptide; Amidation.

FT MOD RES 5 5 AMIDATION.
SQ SEQUENCE 5 AA; 630 MW; 668761F2C9A00000 CRC64;

Query Match 14.3%; Score 1; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 P 5
DB 4 P 4

RESULT 77
E104 LITRU STANDARD; PRT; 5 AA.

AC P82100;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Electrin 4.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Pelodyadinae; Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT "Peptides from the skin glands of the Australian buzzing tree frog Litoria electrica. Comparison with the skin peptides from Litoria rubella.";
RL Aust. J. Chem. 52:639-645 (1999).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Skin.
KW Amphibian defense peptide; Amidation.
FT MOD RES 5 5 AMIDATION.
SQ SEQUENCE 5 AA; 616 MW; 61F2D1A059A00000 CRC64;

Query Match 14.3%; Score 1; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 T 1
DB 3 T 3

RESULT 78
FARP ARTTR STANDARD; PRT; 5 AA.

AC P41853;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE FMRamide-like neuropeptide RYRFP-amide.
OS Artiposthia triangulata (New Zealand flatworm).
OC Eukaryota; Metazoa; Platyhelminthes; Turbellaria; Seriata; Tricladida;
OC Terricola; Geoplanidae; Arthurdendyus.
OX NCBI_TaxID=132421;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RX MEDLINE=94211327; PubMed=7909164;
RA Maule A.G., Shaw C., Halton D.W., Curry W.J., Thim L.;
RT "RYRFPamide: a turbellarian FMRamide-related peptide (FARP).";
RL Regul. Pept. 50:37-43 (1994).
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE) FAMILY.
KW Neuropeptide; Amidation.
FT MOD RES 5 5 AMIDATION.
SQ SEQUENCE 5 AA; 754 MW; 69D4004B44600000 CRC64;

Query Match 14.3%; Score 1; DB 1; Length 5;

Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 Y 6
DB 2 Y 2

RESULT 79
PAP2 PARMA STANDARD; PRT; 5 AA.

AC P81864;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Pardaxin II (PXII) (Fragment).
OS Pardachirus marmoratus (Red sea mores sole).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Soleiidae; Soleidae; Pardachirus.
OX NCBI_TaxID=31087;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RX MEDLINE=87057369; PubMed=3782138;
RA Lazarovici P., Primor N., Loew L.M.;
RT "Purification and pore-forming activity of two hydrophobic polypeptides from the secretion of the Red sea mores sole (Pardachirus marmoratus).";
RL J. Biol. Chem. 261:16704-16713 (1986).
CC -!- FUNCTION: Exhibits unusual shark repellent and surfactant properties. Forms voltage-dependent, ion-permeable channels in membranes. At high concentration causes cell membrane lysis.
CC -!- SUBUNIT: Monomer. In aqueous solution exists as a tetramer.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE PARDAXIN FAMILY.
KW Toxin.
FT NON TER 5 5
SQ SEQUENCE 5 AA; 614 MW; 7769C9C9C8100000 CRC64;

Query Match 14.3%; Score 1; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 G 2
DB 1 G 1

RESULT 80
PRCT PERAM STANDARD; PRT; 5 AA.

AC P01373;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Proctolin.
OS Periplaneta americana (American cockroach).
OS Limulus polyphemus (Atlantic horseshoe crab), and
OC Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blattoidea;
OC Blattidae; Periplaneta.
OX NCBI_TaxID=6978, 6850, 6759;
RN [1]
RP SEQUENCE.
RC SPECIES=P.americana;
RX MEDLINE=76074708; PubMed=576;
RA Starratt A.N., Brown B.E.;
RT "Structure of the pentapeptide proctolin, a proposed neurotransmitter in insects.";
RL Life Sci. 17:1253-1256 (1975).

[2]
 RN BIOLOGICAL SOURCE.
 RP SPECIES=P.americana;
 RC MEDLINE=81225865; PubMed=6113690;
 RA O'Shea M., Adams M.E.;
 RT "Pentapeptide (proctolin) associated with an identified neuron."; Science 213:567-569(1981).
 RL Science 213:567-569(1981).
 RN [3]
 RP SEQUENCE.
 RC SPECIES=L.polyphemus;
 RX MEDLINE=90287800; PubMed=2356151;
 RA Groome J.R., Tillinghast E.K., Townley M.A., Vetrovs A.,
 RA Watson W.H. III, Hunt D.F., Griffin P.R., Alexander J.E.,
 RA Shabanowitz J.;
 RT "Identification of proctolin in the central nervous system of the horseshoe crab, Limulus polyphemus."; Peptides 11:205-211(1990).
 RL Peptides 11:205-211(1990).
 RN [4]
 RP SEQUENCE.
 RC SPECIES=C.maenas;
 RX MEDLINE=86232789; PubMed=2872661;
 RA Stangier J., Dirksen H., Keller R.;
 RT "Identification and immunocytochemical localization of proctolin in pericardial organs of the shore crab, Carcinus maenas."; Peptides 7:67-72(1986).
 RL Peptides 7:67-72(1986).
 CC -!- FUNCTION: STIMULATES CARDIAC OUTPUT AND HINDGUT MOTILITY.
 CC MODULATES VISCERAL AND SKELETAL MUSCLE IN MANY ARTHROPODS.
 CC -!- TISSUE SPECIFICITY: FOUND IN THE LATERAL WHITE NEURONS AND IN THE CRAB PERICARDIAL ORGANS.
 CC PIR; A01644; HOROHA.
 DR PIR; A60411; A60411.
 KW Neuropeptide.
 SQ SEQUENCE 5 AA; 649 MW; 71B7673B4460000 CRC64;
 Query Match 14.3%; Score 1; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 6 Y 6
 Db 2 Y 2
 RESULT 81
 PSK DAUCA STANDARD; PRT; 5 AA.
 AC P50261;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Phytosulfokine-alpha (PSK-alpha) [Contains: Phytosulfokine-beta (PSK-beta)].
 DE Daucus carota (Carrot).
 OS Daucus carota (Carrot).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; campanulids; Apiales; Apiaceae; Daucus.
 OX NCBI_TaxID=4039;
 RN [1]
 RP SEQUENCE, AND IDENTIFICATION BY MASS SPECTROMETRY.
 RC STRAIN=cv. US-Harumakigosun;
 RX MEDLINE=2012743; PubMed=10750705;
 RA Hanai H., Matsuno T., Yamamoto M., Matsubayashi Y., Kobayashi T., Kamada H., Sakagami Y.;
 RA "A secreted peptide growth factor, phytosulfokine, acting as a stimulatory factor of carrot somatic embryo formation."; Plant Cell Physiol. 41:27-32(2000).
 RL Plant Cell Physiol. 41:27-32(2000).
 CC -!- FUNCTION: IN PRESENCE OF 2,4-D, STIMULATES PROLIFERATION OF THE CELLS, BUT DOES NOT STIMULATE DIFFERENTIATION INTO THE SOMATIC EMBRYOS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- PFM: SULFATION IS IMPORTANT FOR ACTIVITY AND FOR THE BINDING TO A PUTATIVE MEMBRANE RECEPTOR (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE PHYTOSULFOKINE FAMILY.

KW Growth factor; Sulfation.
 FT PEPTIDE 1 4 PHYTOSULFOKINE-BETA.
 FT MOD_RES 1 1 SULFATION.
 FT MOD_RES 3 3 SULFATION.
 SQ SEQUENCE 5 AA; 687 MW; 76C1BB504B300000 CRC64;
 Query Match 14.3%; Score 1; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 6 Y 6
 Db 1 Y 1
 RESULT 82
 RE21_LITRU STANDARD; PRT; 5 AA.
 AC P82071;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Rubellidin 2.1.
 OS Litoria rubella (Desert tree frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
 OC Pelodyadinae; Litoria.
 OX NCBI_TaxID=104895;
 RN [1]
 RP SEQUENCE, AND MASS SPECTROMETRY.
 RC TISSUE=Skin secretion;
 RA Steinbörner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C., Tyler M.J., Wallace J.C.;
 RT "The structure of new peptides from the Australian red tree frog 'Litoria rubella'. The skin peptide profile as a probe for the study of evolutionary trends of amphibians."; Aust. J. Chem. 49:955-963(1996).
 RL Aust. J. Chem. 49:955-963(1996).
 CC -!- FUNCTION: Shows neither neuropeptide activity nor antibiotic activity.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.
 CC -!- MASS SPECTROMETRY: MW=626; METHOD=FAB.
 KW Amphibian defense peptide.
 SQ SEQUENCE 5 AA; 626 MW; 6DD9C9CB10300000 CRC64;
 Query Match 14.3%; Score 1; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 3 E 3
 Db 2 E 2
 RESULT 83
 RE31_LITRU STANDARD; PRT; 5 AA.
 AC P82072;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Rubellidin 3.1.
 OS Litoria rubella (Desert tree frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
 OC Pelodyadinae; Litoria.
 OX NCBI_TaxID=104895;
 RN [1]
 RP SEQUENCE, AND MASS SPECTROMETRY.
 RC TISSUE=Skin secretion;
 RA Steinbörner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C., Tyler M.J., Wallace J.C.;
 RT "The structure of new peptides from the Australian red tree frog

RT 'Litoria rubella'. The skin peptide profile as a probe for the study
 of evolutionary trends of amphibians.";

RL Aust. J. Chem. 49:955-963(1996).

CC -!- FUNCTION: Shows neither neuropeptide activity nor antibiotic
 activity.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.

CC -!- MASS SPECTROMETRY: MW=655; METHOD=FAB.

KW Amphibian defense peptide; Amidation.

FT MOD RES 5 5 AMIDATION.

SQ SEQUENCE 5 AA; 656 MW; 71A9C9CB10300000 CRC64;

Query Match 14.3%; Score 1; DB 1; Length 5;

Best Local Similarity 100.0%; Pred. No. 1.3e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 E 3

Db 2 E 2

RESULT 84

ID RE32 LITRU STANDARD; PRT; 5 AA.

AC P82073;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Rubellidin 3.2.

OS Litoria rubella (Desert tree frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;

OC Pelodyadinae; Litoria.

OX NCBI_TaxID=104895;

RN [1]

RP SEQUENCE.

RC TISSUE-Skin secretion;

RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;

RT "Peptides from the skin glands of the Australian buzzing tree frog

Litoria electrica. Comparison with the skin peptides from Litoria

rubella.";

RL Aust. J. Chem. 52:639-645(1999).

CC -!- FUNCTION: Shows neither neuropeptide activity nor antibiotic

activity.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.

KW Amphibian defense peptide.

SQ SEQUENCE 5 AA; 570 MW; 71A9C9C862A00000 CRC64;

Query Match 14.3%; Score 1; DB 1; Length 5;

Best Local Similarity 100.0%; Pred. No. 1.3e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 G 2

Db 2 G 2

RESULT 85

ID SUGA ACHDO STANDARD; PRT; 5 AA.

AC P19991;

DT 01-FEB-1991 (Rel. 17, Created)

DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE Subesophageal ganglion pentapeptide.

OS Acheta domestica (House cricket).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Orthoptera; Orthopteroidea; Orthoptera; Ensifera; Gryllidae; Gryllinae;

OC Acheta.

OX NCBI_TaxID=6997;

RN [1]

RP SEQUENCE.

RA Wicker C., Wicker C.;

RT "Isolation and structure of a peptide isolated from the

subesophageal ganglion of Acheta domestica (Orthoptera).";

RL Comp. Biochem. Physiol. 88C:185-187(1987).

CC -!- SUBCELLULAR LOCATION: MAIN PEPTIDE FROM THE SUBESOPHAGEAL

CC GANGLIA.

DR PIR; JS0319; JS0319.

SQ SEQUENCE 5 AA; 476 MW; 69D76DDDDDB00000 CRC64;

Query Match 14.3%; Score 1; DB 1; Length 5;

Best Local Similarity 100.0%; Pred. No. 1.3e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 P 5

Db 4 P 4

RESULT 86

TPIS CANFA

ID TPIS CANFA STANDARD; PRT; 5 AA.

AC P54714;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE Triosephosphate isomerase (EC 5.3.1.1) (TIM) (Fragment).

GN TP11.

OS Canis familiaris (Dog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

OX NCBI_TaxID=9615;

RN [1]

RP SEQUENCE.

RC TISSUE=Heart;

RX MEDLINE=98163340; PubMed=9504812;

RA Dunn M.J., Corbett J.M., Wheeler C.H.;

RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of

dog heart proteins.";

RL Electrophoresis 18:2795-2803(1997).

CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate = glyceralone

phosphate.

CC -!- PATHWAY: Plays an important role in several metabolic pathways.

CC -!- SUBUNIT: Homodimer (By similarity).

CC -!- SIMILARITY: BELONGS TO THE TRIOSEPHOSPHATE ISOMERASE FAMILY.

DR HSC-2DPAGE; P54714; DOG.

DR InterPro; IPR000652; Triophos ismrse.

DR PROSITE; PS00171; TIM; PARTIAL.

KW Isomerase; Glycolysis; Gluconeogenesis; Fatty acid biosynthesis;

KM Pentose shunt.

FT NON_TER 1 1

FT NON_TER 5 5

SQ SEQUENCE 5 AA; 550 MW; 64444862C9A00000 CRC64;

Query Match 14.3%; Score 1; DB 1; Length 5;

Best Local Similarity 100.0%; Pred. No. 1.3e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 G 2

Db 3 G 3

RESULT 87

TRM3 ECOLI

ID TRM3 ECOLI STANDARD; PRT; 5 AA.

AC P13973;

DT 01-JAN-1990 (Rel. 13, Created)

DT 01-JAN-1990 (Rel. 13, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE TrmM protein (Fragment).

GN TRAM.

OS Escherichia coli.

OG Plasmid IncFII R100.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 [1]
 RN SEQUENCE FROM N.A. PubMed=2836369;
 RA MEDLINE=86227859; Ohtsubo B.;
 RA Inamoto S., Yoshioke Y., Ohtsubo B.;
 RT "Identification and characterization of the products from the traJ
 RT and traY genes of plasmid R100.";
 RL J. Bacteriol. 170:2749-2757(1988).
 CC -!- FUNCTION: TRANSFER GENE PROTEIN. IS INVOLVED IN THE CONJUGATION
 CC OF BACTERIAL CELLS FOR THE EXCHANGE OF PLASMID DNA.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: BELONGS TO THE TRAM FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M20941; ; NOT_ANNOTATED_CDS.
 DR PIR; A32014; A32014.
 KW Conjugation; Plasmid; DNA-binding.
 FT NON_TER 1
 FT 1
 SQ SEQUENCE 5 AA; 634 MW; 6B1B1AA443500000 CRC64;

Query Match 14.3%; Score 1; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 K 4
 Db 1 K 1

RESULT 88
 UC22 MAIZE
 ID UC22 MAIZE STANDARD; PRT; 5 AA.
 AC P80628;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Unknown protein from 2D-page of etiolated coleoptile (Spot 474)
 DE (Fragment).
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Coleoptile;
 RA Touzet P., Ricciardi F., Morin C., Damerl C., Huet J.-C.,
 RA Pernollet J.-C., Zivy M., de Vienne D.;
 RT "The maize two dimensional gel protein database: towards an integrated
 RT genome analysis program.";
 RL Theor. Appl. Genet. 93:997-1005(1996).
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 6.1, ITS MW IS: 30.4 kDa.
 DR Maize-2DPAGE; P80628; COLEOPTILE.
 DR MaizeDB; 123954; --
 FT NON_TER 1
 FT 1
 SQ SEQUENCE 5 AA; 654 MW; 72CB19C9C03000000 CRC64;

Query Match 14.3%; Score 1; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 E 3

Db 4 E 4
 RESULT 89
 UF01 MOUSE
 ID UF01 MOUSE STANDARD; PRT; 5 AA.
 AC P38639;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE Unknown protein from 2D-page of fibroblasts (P19) (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Fibroblast;
 RX MEDLINE=95009907; PubMed=7523108;
 RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
 RT "Separation and sequencing of familial and novel murine proteins
 RT using preparative two-dimensional gel electrophoresis.";
 RL Electrophoresis 15:735-745(1994).
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 6.6, ITS MW IS: 19 kDa.
 FT NON_TER 5
 FT 5
 SQ SEQUENCE 5 AA; 717 MW; 7364087043100000 CRC64;

Query Match 14.3%; Score 1; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 G 2
 Db 3 G 3

RESULT 90
 UX44 CHLTR
 ID UX44 CHLTR STANDARD; PRT; 5 AA.
 AC P38005;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Unknown protein from 2D-page from elementary body (Fragment).
 DE Chlamydia trachomatis.
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=813;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=L2/434/Bu;
 RA Bini L., Santucci A., Magi B., Marzocchi B., Sanchez-Campillo M.,
 RA Comanducci M., Christian G., Birkelund S., Vretou E., Ratti G.,
 RA Pallini V.;
 RT Submitted (SEP-1994) to the SWISS-PROT data bank.
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 4.5, ITS MW IS: 28 kDa.
 DR Siena-2DPAGE; P38005; --
 FT NON_TER 5
 FT 5
 SQ SEQUENCE 5 AA; 474 MW; 75BAA865AA8000000 CRC64;

Query Match 14.3%; Score 1; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 G 2
 Db 3 G 3

RESULT 91
 ACPH_RABIT

ID ACPH RABIT STANDARD; PRT; 6 AA.
 AC P25154;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Acylamino-acid-releasing enzyme (EC 3.4.19.1) (Acyl-peptide hydrolase)
 DE (APH) (Acylaminoacyl-peptidase) (Fragment).
 GN APEH.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP TISSUE=Muscle;
 RC MEDLINE=92222120; PubMed=1807161;
 RA Krishna R.G., Chin C.C.Q., Wold F.;
 RT "N-terminal sequence analysis of N alpha-acetylated proteins after
 RT unblocking with N-acylaminoacyl-peptide hydrolase.";
 RL Anal. Biochem. 199;45-50(1991).
 CC -!- FUNCTION: THIS ENZYME CATALYZES THE HYDROLYSIS OF THE AMINO-
 CC TERMINAL PEPTIDE BOND OF AN N-ACETYLATED PEPTIDE TO GENERATE
 CC AN N-ACETYLATED AA AND A PEPTIDE WITH A FREE AMINO-TERMINUS.
 CC IT PREFERENTIALLY CLEAVES OFF AC-ALA, AC-MET AND AC-SER.
 CC -!- CATALYTIC ACTIVITY: Acylaminoacyl-peptide + H(2)O = acylamino acid
 CC + peptide.
 CC -!- SUBUNIT: Homotetramer.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S9C.
 DR PIR; A49792; A49792.
 DR MEROPS; S09.004; -.
 DR InterPro; IPR002471; Prol endopep ser.
 DR PROSITE; PS00708; PRO_ENDOPEP_SER; PARTIAL.
 KW Hydrolase; Acetylation.
 FT MOD_RES 1 1 ACETYLATION.
 FT NON_TER 6
 SQ SEQUENCE 6 AA; 775 MW; 6732D6C40B16F000 CRC64;

 Query Match 14.3%; Score 1; DB 1; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 3 E 3
 Db 2 E 2

 RESULT 92
 ASP2 LACSN STANDARD; PRT; 6 AA.
 AC P82655;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Acid shock protein 2 (fragment).
 OS Lactobacillus sanfranciscensis (Lactobacillus sanfrancisco).
 OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
 OC Lactobacillus.
 OX NCBI_TaxID=1625;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=CB1;
 RX MEDLINE=21322712; PubMed=11429463;
 RA De Angelis M., Bini L., Pallini V., Cocconcelli P.S., Gobetti M.;
 RT "The acid-stress response in Lactobacillus sanfranciscensis CB1.";
 RL Microbiology 147;1863-1873(2001).
 CC -!- INDUCTION: Overexpressed in acid environments.
 FT NON_TER 6
 SQ SEQUENCE 6 AA; 778 MW; 6AA45B5B132A8000 CRC64;

 Query Match 14.3%; Score 1; DB 1; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 K 4
 Db 2 K 2

 RESULT 93
 CIPI MYTED STANDARD; PRT; 6 AA.
 ID CIPI MYTED
 AC P13736;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Contraction-inhibiting peptide I (MIP I).
 OS Mytilus edulis (Blue mussel).
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
 OC Mytiloidea; Mytilidae; Mytilus.
 OX NCBI_TaxID=6550;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Pedal ganglion;
 RX MEDLINE=88240357; PubMed=3377776;
 RA Hirata T., Kubota I., Iwasawa N., Takabatake I., Ikeda T., Muneoka Y.;
 RT "Structures and actions of Mytilus inhibitory peptides.";
 RL Biochem. Biophys. Res. Commun. 152;1376-1382(1988).
 CC -!- FUNCTION: INHIBITORY ACTION ON CONTRACTIONS IN SEVERAL MOLLUSCAN
 CC MUSCLES.
 CC -!- SIMILARITY: TO MIP II.
 DR PIR; A27696; A27696.
 KM Hormone; Amidation.
 FT MOD_RES 6
 SQ SEQUENCE 6 AA; 637 MW; 72C9C6875B81000 CRC64;

 Query Match 14.3%; Score 1; DB 1; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 2 G 2
 Db 1 G 1

 RESULT 94
 CIP2 MYTED STANDARD; PRT; 6 AA.
 ID CIP2 MYTED
 AC P13737;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Contraction-inhibiting peptide II (MIP II).
 OS Mytilus edulis (Blue mussel).
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
 OC Mytiloidea; Mytilidae; Mytilus.
 OX NCBI_TaxID=6550;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Pedal ganglion;
 RX MEDLINE=88240357; PubMed=3377776;
 RA Hirata T., Kubota I., Iwasawa N., Takabatake I., Ikeda T., Muneoka Y.;
 RT "Structures and actions of Mytilus inhibitory peptides.";
 RL Biochem. Biophys. Res. Commun. 152;1376-1382(1988).
 CC -!- FUNCTION: INHIBITORY ACTION ON CONTRACTIONS IN SEVERAL MOLLUSCAN
 CC MUSCLES.
 CC -!- SIMILARITY: TO MIP I.
 DR PIR; B27696; B27696.
 KM Hormone; Amidation.
 FT MOD_RES 6
 SQ SEQUENCE 6 AA; 621 MW; 72C9C6876DD81000 CRC64;

 Query Match 14.3%; Score 1; DB 1; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 G 2

Db 1 G 1

RESULT 95

E101 LITRU STANDARD; PRT; 6 AA.
 ID AC P82056;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Electrin 1.
 OS Litoria rubella (Desert tree frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
 OC Pelodyadinae; Litoria.
 OX NCBI_TaxID=104895;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Skin secretion;
 RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
 RT "Peptides from the skin glands of the Australian buzzing tree frog
 RT Litoria electrica. Comparison with the skin peptides from Litoria
 RT rubella.";
 RL Aust. J. Chem. 52:639-645(1999).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Skin.
 KW Amphibian defense peptide; Amidation.
 FT MOD.RES 6 6
 SQ SEQUENCE 6 AA; 792 MW; 6683704772C9A000 CRC64;

Query Match 14.3%; Score 1; DB 1; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 P 5

Db 3 P 3

RESULT 96

FARP_MONEX STANDARD; PRT; 6 AA.
 ID AC P41966;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE FMRPamide-like neuropeptide GNFRP-amide.
 OS Moniezia expansa (Sheep tapeworm).
 OC Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
 OC Cyclophyllidae; Anoplocephalidae; Moniezia.
 OX NCBI_TaxID=28841;
 RN [1]
 RP SEQUENCE.
 RC MEDLINE=93312289; PubMed=8323531;
 RA Maule A.G., Shaw C., Halton D.W., Thim L.;
 RT "GNFRPamide: a novel FMRPamide-immunoreactive peptide isolated from
 RT the sheep tapeworm, Moniezia expansa.";
 RL Biochem Biophys Res Commun. 193:1054-1060(1993).
 CC -!- SIMILARITY: BELONGS TO THE FARP (FMRPAMIDE RELATED PEPTIDE)
 CC FAMILY.
 DR PIR; A43129; A43129.
 KW Neuropeptide; Amidation.
 FT MOD.RES 6 6
 SQ SEQUENCE 6 AA; 787 MW; 69D409C9C4481000 CRC64;

Query Match 14.3%; Score 1; DB 1; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 G 2

Db 1 G 1

RESULT 97

LOKJ_LOCKI STANDARD; PRT; 6 AA.
 ID AC P41491;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Locustakinin 1.
 OS Locusta migratoria (Migratory locust).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;
 OC Acridoidea; Acrididae; Oedipodinae; Locusta.
 OX NCBI_TaxID=7004;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Corpora cardiaca;
 RC MEDLINE=92262851; PubMed=1585017;
 RA Schoofs L., Holman G.M., Proost P., van Damme J., Hayes T.K.,
 RA de Loof A.;
 RT "Locustakinin, a novel myotropic peptide from Locusta migratoria,
 RT isolation, primary structure and synthesis.";
 RL Regul. Pept. 37:49-57(1992).
 CC -!- FUNCTION: Myotropic peptide. May be important in the stimulation
 CC of ion transport and inhibition of diuretic activity in Malpighian
 CC tubules.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC PIR; A61068; A61068.
 DR Neuropeptide; Amidation.
 FT MOD.RES 6 6
 SQ SEQUENCE 6 AA; 654 MW; 686365A5B9CDB000 CRC64;

Query Match 14.3%; Score 1; DB 1; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 G 2

Db 6 G 6

RESULT 98

TMOF_SARBU STANDARD; PRT; 6 AA.
 ID AC P41495;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Trypsin-modulating oostatic factor (TMOP).
 OS Sarcophaga bullata (Grey flesh fly) (Neobellieria bullata).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
 OC Sarcophagidae; Sarcophaga.
 OX NCBI_TaxID=7385;
 RN [1]
 RP SEQUENCE, AND SYNTHESIS.
 RC TISSUE=Ovary;
 RX MEDLINE=94211930; PubMed=8159807;
 RA Bylemans D., Borovsky D., Hunt D.F., Shabanowitz J., Grauwels L.,
 RA de Loof A.;
 RT "Sequencing and characterization of trypsin modulating oostatic
 RT factor (TMOF) from the ovaries of the grey fleshfly, Neobellieria
 RT (Sarcophaga) bullata.";
 RL Regul. Pept. 50:61-72(1994).
 CC -!- FUNCTION: HAS AN OOSTATIC ACTIVITY. INHIBITS TRYPSIN BIOSYNTHESIS
 CC IN THE MIDGUT WHICH INDIRECTLY REDUCES THE VITELLOGENIN
 CC CONCENTRATION IN THE HEMOLYPH RESULTING IN INHIBITION OF OOCYTE
 CC DEVELOPMENT.
 CC -!- DEVELOPMENTAL STAGE: SYNTHESIZED AND RELEASED FROM FOLLICULAR
 CC EPITHELIUM AFTER A BLOOD MEAL.
 KW Hormone.
 SQ SEQUENCE 6 AA; 695 MW; 61E72451B7642000 CRC64;

```

Query Match      14.3%; Score 1; DB 1; Length 6;
Best Local Similarity 100.0%; Pred.No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 P 5
DB      2 P 2

RESULT 99
TRPI_PSEPU
ID TRPI_PSEPU STANDARD; PRT; 6 AA.
AC P36414;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE TRPB operon transcriptional activator (fragment).
GN TRPI.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PPG1 CIS;
RX MEDLINE=89335826; PubMed=2503057;
RA Eberly L., Crawford I.P.;
RT "DNA sequence of the tryptophan synthase genes of Pseudomonas putida.";
RL Biochimie 71:521-531(1989).
CC -!- FUNCTION: ACTIVATES THE EXPRESSION OF THE TRPB GENES ENCODING THE TWO TRYPTOPHAN SYNTHASE SUBUNITS. IN THE ABSENCE OF THE INDUCER (INDOLEGLYCEROL PHOSPHATE), TRPI BINDS UPSTREAM OF THE TRPB OPERON, OVERLAPPING ITS OWN PROMOTER REGION.
CC -!- SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL REGULATORS.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
EMBL; X13299; CAA31660.1; -.
InterPro; IPR000847; HTH_LYSR.
PROSITE; PS00044; HTH_LYSR_FAMILY; PARTIAL.
KW Tryptophan biosynthesis; Transcription regulation; Activator;
KW DNA-binding.
FT NON TER
FT NON TER
SQ SEQUENCE 6 AA; 683 MW; 77672AA1EDD6F000 CRC64;

```

```

Query Match      14.3%; Score 1; DB 1; Length 6;
Best Local Similarity 100.0%; Pred.No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      5 P 5
DB      6 P 6

```

```

RESULT 100
UN06_CLOPA
ID UN06_CLOPA STANDARD; PRT; 6 AA.
AC P81351;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Unknown protein CP 6 from 2D-page (fragment).
OS Clostridium pasteurianum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

```

```

OC Clostridium.
OX NCBI_TaxID=1501;
RN [1]
RP SEQUENCE.
RC STRAIN=W5;
RX MEDLINE=98291870; PubMed=9629918;
RA Flengsrud R., Skjeldal L.;
RT "Two-dimensional gel electrophoresis separation and N-terminal sequence analysis of proteins from Clostridium pasteurianum W5.";
RL Electrophoresis 19:802-806(1998).
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN PROTEIN IS: 5.0, ITS MW IS: 75.9 kDa.
FT NON TER
FT NON TER
SQ SEQUENCE 6 AA; 657 MW; 605B1DC1A45A8000 CRC64;

Query Match      14.3%; Score 1; DB 1; Length 6;
Best Local Similarity 100.0%; Pred.No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 T 1
DB      3 T 3

```

Search completed: October 21, 2003, 18:24:28
Job time : 18.0556 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:24:06 ; Search time 22.5556 Seconds
(without alignments)
29.845 Million cell updates/sec

Title: US-10-057-890A-8

Perfect score: 7

Sequence: 1 TGEKPYK 7

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1100

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Listing first 150 summaries

Database : PIR 76.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3	42.9	6	2	PT0516
2	3	42.9	6	2	PT0662
3	3	42.9	8	2	I48935
4	3	42.9	9	2	D48186
5	3	42.9	10	2	A42089
6	2	28.6	3	3	PT0636
7	2	28.6	4	2	A02147
8	2	28.6	4	2	E44823
9	2	28.6	4	2	PT0677
10	2	28.6	5	2	D60274
11	2	28.6	5	2	B45525
12	2	28.6	5	2	C53284
13	2	28.6	5	2	PT0686
14	2	28.6	5	2	PT0540
15	2	28.6	5	2	PT0590
16	2	28.6	6	2	S11556
17	2	28.6	6	2	B44510
18	2	28.6	6	2	S14159
19	2	28.6	6	2	B61512
20	2	28.6	6	2	A61140
21	2	28.6	6	2	S78764
22	2	28.6	6	2	A20186
23	2	28.6	6	2	B33932
24	2	28.6	6	2	PT0518
25	2	28.6	6	2	PT0630
26	2	28.6	6	2	PT0643
27	2	28.6	6	2	PT0619
28	2	28.6	6	2	PT0668
29	2	28.6	6	2	PT0723

T-cell receptor be
hypothalamic hepta
peptidyl-dipectida
opacity protein p.
cryptophyllin, bas
membrane protein -
microcin C7 - Esch
18K protein 5507 -
T-cell receptor be
T-cell receptor be
T-cell receptor be
T-cell receptor be
T-cell receptor be
T-cell receptor be
hypothetical pepti
adipokinetic hormo
leghemoglobin iii
rpsA protein - Erw
disimilatory sulf
glycine reductase
nitrate reductase
hypothetical prote
unidentified 6.5/3
inulinase (EC 3.2.
variant surface gl
neuropeptide calla
serum albumin - do
apolipoprotein A-I
olfactory receptor
T-cell receptor be
T-cell receptor be
T-cell receptor be
neural cell adheci
aspartate kinase (
litorin - Rohde's
phylocaerulein -
coat protein beta
delta sleep-induci
quinoline 2-oxidor
disimilatory sulf
cell surface adhes
tetrameric protein
protein QA30033 -
endosperm protein,
hypothetical prote
oxytocin-related p
cardioactive pepti
cardioactive pepti
cardioactive pepti
calliFMRFamide 4 -
cytochrome-c oxida
3', 5'-cyclic-GMP p
MHC class I histoc
118K stomach cance
Ig heavy chain CRD
Ig heavy chain CRD
macrophage inhibit
zymogen granule me
cyokeratin 4 - bo
enamelin 1 - bovin
Ig H chain V-D-J r
late GI-69 protein
bone gla protein -
T-cell receptor be
T-cell receptor be

```

103 2 28.6 9 2 PH0918
104 2 28.6 9 2 S39767
105 2 28.6 10 1 RHLWGS
106 2 28.6 10 2 A43405
107 2 28.6 10 2 A61289
108 2 28.6 10 2 A61337
109 2 28.6 10 2 H60787
110 2 28.6 10 2 D60788
111 2 28.6 10 2 B60787
112 2 28.6 10 2 A60788
113 2 28.6 10 2 C60589
114 2 28.6 10 2 I60588
115 2 28.6 10 2 H60588
116 2 28.6 10 2 D60588
117 2 28.6 10 2 A49581
118 2 28.6 10 2 B49581
119 2 28.6 10 2 B61033
120 2 28.6 10 2 B33143
121 2 28.6 10 2 A33143
122 2 28.6 10 2 C35389
123 2 28.6 10 2 A44646
124 2 28.6 10 2 I44644
125 2 28.6 10 2 JP0072
126 2 28.6 10 2 PT0084
127 2 28.6 10 2 S62880
128 2 28.6 10 2 E61512
129 2 28.6 10 2 H61308
130 2 28.6 10 2 A45187
131 2 28.6 10 2 B56899
132 2 28.6 10 2 A24196
133 2 28.6 10 2 B45482
134 2 28.6 10 2 S23371
135 2 28.6 10 2 F49033
136 2 28.6 10 2 S71948
137 2 28.6 10 2 PH1633
138 2 28.6 10 2 S65385
139 2 28.6 10 2 PH0927
140 2 28.6 10 2 PH0946
141 2 28.6 10 2 PH0925
142 2 28.6 10 2 PH0926
143 2 28.6 10 2 AS3173
144 2 28.6 10 2 PA0116
145 2 28.6 10 2 PS0220
146 1 14.3 3 3 RHTD70
147 1 14.3 3 3 RHPGT
148 1 14.3 3 3 RSHST
149 1 14.3 3 3 A92971
150 1 14.3 3 3 GRHU

ALIGNMENTS

RESULT 1
PT0516
T-cell receptor beta chain V-D-J region (100-4AF) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0516
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A>Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0516
A>Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-6 <FE>
A:Experimental source: adult thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match 42.9%; Score 3; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 1
PT0516
T-cell receptor beta chain V-D-J region (100-4AF) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0516
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A>Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0516
A>Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-6 <FE>
A:Experimental source: adult thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match 42.9%; Score 3; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 GEK 4
DB 3 GEK 5

RESULT 2
PT0662
T-cell receptor beta chain V-D-J region (121-3BA) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0662
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A>Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0662
A>Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-6 <FE>
A:Experimental source: day 4 postnatal thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match 42.9%; Score 3; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGE 3
DB 4 TGE 6

RESULT 3
I48935
apolipoprotein A-II - western wild mouse (fragment)
C:Species: Mus spretus (western wild mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C:Accession: I48935
R:Ko, M.S.; Wang, X.; Horton, J.H.; Hagen, M.D.; Takahashi, N.; Maezaki, Y.; Nadeau, J.J.
Mamm. Genome 5, 349-355, 1994
A>Title: Genetic mapping of 40 cDNA clones on the mouse genome by PCR.
A:Reference number: I48934; MUID:94319082; PMID:8043949
A:Accession: I48935
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-8 <RES>
A:Cross-references: EMBL:U05692; NID:9497011; PIDN:AAB60463.1; PID:G642827

Query Match 42.9%; Score 3; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EXP 5
DB 1 EXP 3

RESULT 4
D48186
ATPase R1 subunit - wood tobacco (fragment)
C:Species: Nicotiana sylvestris (wood tobacco)
C>Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 23-Feb-1997
C:Accession: D48186
R:De Paeppe, R.; Forchioni, A.; Chetrit, P.; Vedel, F.
Proc. Natl. Acad. Sci. U.S.A. 90, 5934-5938, 1993
A>Title: Specific mitochondrial proteins in pollen: presence of an additional ATP synth.
A:Reference number: A48186; MUID:93317598; PMID:8327463
A:Accession: D48186
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-9 <DEI>
A:Experimental source: pollen
A>Note: sequence extracted from NCBI backbone (NCBI:134871)

```


Query Match 28.6%; Score 2; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GE 3
||
3 GE 4

Db

RESULT 10
D60274
major protein antigen MPT46 - Mycobacterium tuberculosis (fragment)
C:Species: Mycobacterium tuberculosis
C:Date: 11-Dec-1992 #sequence_revision 11-Dec-1992 #text_change 30-Sep-1993
C:Accession: D60274
R:Nagai, S.; Wiker, H.G.; Harboe, M.; Kinomoto, M.
Infect. Immun. 59, 372-382, 1991
A:Title: Isolation and partial characterization of major protein antigens in the culture
A:Reference number: A60274; MUID:9109989; PMID:1898899
A:Accession: D60274
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-5 <NAG>

Query Match 28.6%; Score 2; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EK 4
||
4 EK 5

Db

RESULT 11
B45525
actin 1 - malaria parasite (Plasmodium falciparum) (fragments)
C:Species: Plasmodium falciparum
C:Date: 03-Jun-1993 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000
C:Accession: B45525
R:Wesseling, J.G.; Snijders, P.J.F.; van Someren, P.; Jansen, J.; Smits, M.A.; Schoenmak
Mol. Biochem. Parasitol. 35, 167-176, 1989
A:Title: Stage-specific expression and genomic organization of the actin genes of the ma
A:Reference number: A45525; MUID:89364996; PMID:2671721
A:Accession: B45525
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-5 <WES>
A:Cross-references: GB:J03988
A:Note: the authors translated the codon GAA for residue 3 as Gly
C:Comment: The actin 1 gene contains no introns.

Query Match 28.6%; Score 2; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GE 3
||
2 GE 3

Db

RESULT 12
C53284
T-cell receptor beta 2 chain D region, Dbeta2 - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 02-May-1994 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C:Accession: C53284
R:Harindranath, N.; Alexander, C.B.; Mage, R.G.
Mol. Immunol. 28, 881-888, 1991
A:Title: Evolutionarily conserved organization and sequences of germline diversity and j
A:Reference number: A53284; MUID:91342695; PMID:1678859
A:Accession: C53284
A:Status: preliminary
A:Molecule type: DNA

Query Match 28.6%; Score 2; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GE 3
||
3 GE 4

Db

RESULT 15
PT0590

A:Residues: 1-5 <HAR>
A:Cross-references: GB:S60737; MUID:9233916; PIDN:AA819519.1; PID:9233919
A:Note: sequence extracted from NCBI backbone (NCBIN:60737, NCBIIP:60740)
C:Keywords: T-cell receptor

Query Match 28.6%; Score 2; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
||
2 TG 3

Db

RESULT 13
PT0686
T-cell receptor beta chain V-D-J region (120-1R and 154-2D) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 05-Sep-1997
C:Accession: PT0615; PT0686
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0615
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-5 <FEE>
A:Experimental source: newborn thymus, strain BALB/c, clone 120-1R
A:Accession: PT0686
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-5 <FE2>
A:Experimental source: day 18 fetal thymus, strain BALB/c, clone 154-2D
C:Keywords: T-cell receptor

Query Match 28.6%; Score 2; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
||
4 TG 5

Db

RESULT 14
PT0540
T-cell receptor beta chain V-D-J region (126-1L) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0540
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0540
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-5 <FEE>
A:Experimental source: day 18 fetal thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match 28.6%; Score 2; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GE 3
||
3 GE 4

Db

RESULT 15
PT0590

T-cell receptor beta chain V-D-J region (141-1B) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 C:Accession: P70590
 E:Feeney, A.J.
 J. Exp. Med. 174, 115-124, 1991
 A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
 A:Reference number: P70509; MUID:91277601; PMID:1711558
 A:Accession: P70590
 A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-5 <FE>
 A:Experimental source: day 19 fetal thymus, strain BAUB/c
 C:Keywords: T-cell receptor

Query Match 28.6%; Score 2; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TG 2
 ||
 Db 4 TG 5

RESULT 16

S11556
 hydrogen sulfite reductase (EC 1.8.99.3) chain 2 - Desulfovibrio thermophilus (fragment)
 N:Alternate names: Bisulfite reductase; desulfotuscin
 C:Species: Desulfovibrio thermophilus
 C:Date: 19-Mar-1997 #sequence_revision 12-Dec-1997 #text_change 30-Jan-1998
 C:Accession: S11556
 R:Fauque, G.; Lino, A.R.; Czechowski, M.; Kang, L.; DerVartanian, D.V.; Moura, J.J.G.; I
 Biochim. Biophys. Acta 1040, 112-118, 1990
 A:Title: Purification and characterization of bisulfite reductase (desulfotuscin) from
 A:Reference number: S11024; MUID:90335276; PMID:2165817
 A:Accession: S11556
 A:Molecule type: protein
 A:Residues: 1-6 <FAU>
 C:Keywords: oxidoreductase

Query Match 28.6%; Score 2; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EK 4
 ||
 Db 3 EK 4

RESULT 17

B44510
 hypothetical protein C (mleR 3' region) - Lactococcus lactis (fragment)
 C:Species: Lactococcus lactis
 C:Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 30-Sep-1993
 C:Accession: B44510
 R:Renault, P.; Gaillardin, C.; Heslot, H.
 J. Bacteriol. 171, 3108-3114, 1989
 A:Title: Product of the Lactococcus lactis gene required for malolactic fermentation is
 A:Reference number: A44510; MUID:89255069; PMID:2498286
 A:Accession: B44510
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-6 <REN>
 A:Cross-references: EMBL:M90762

Query Match 28.6%; Score 2; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EK 4
 ||
 Db 2 EK 3

RESULT 18

S14159
 parasporal crystal protein CryIC - Bacillus thuringiensis (fragment)
 N:Alternate names: delta-endotoxin
 C:Species: Bacillus thuringiensis
 C:Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 18-Jul-1997
 C:Accession: S14159
 R:Conventes, D.; Cherlet, M.; van Damme, J.; Lasters, I.; Lauwereys, M.
 Eur. J. Biochem. 195, 631-635, 1991
 A:Title: Two structural domains as a general fold of the toxic fragment of the Bacillu
 A:Reference number: S14087; MUID:91153300; PMID:1847865
 A:Accession: S14159
 A:Molecule type: protein
 A:Residues: 1-6 <CON>

Query Match 28.6%; Score 2; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TG 2
 ||
 Db 3 TG 4

RESULT 19

B61512
 variant surface glycoprotein MITat 1.1 - Trypanosoma brucei (fragment)
 C:Species: Trypanosoma brucei
 C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 07-May-1999
 C:Accession: B61512
 R:Holder, A.A.; Cross, G.A.M.
 Mol. Biochem. Parasitol. 2, 135-150, 1981
 A:Title: Glycopeptides from variant surface glycoproteins of Trypanosoma brucei. C-ter
 A:Reference number: A61512; MUID:81172836; PMID:6163983
 A:Accession: B61512
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-6 <HOL>
 C:Keywords: glycoprotein

Query Match 28.6%; Score 2; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TG 2
 ||
 Db 4 TG 5

RESULT 20

A61140
 sperm acrosomal protein - spoonworm (Urechis caupo) (fragment)
 C:Species: Urechis caupo
 C:Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 07-Apr-1994
 C:Accession: A61140
 R:Gould, M.C.; Stephano, J.L.
 Dev. Biol. 146, 509-518, 1991
 A:Title: Peptides from sperm acrosomal protein that initiate egg development.
 A:Reference number: A61140; MUID:91232672; PMID:1864468
 A:Accession: A61140
 A:Molecule type: protein
 A:Residues: 1-6 <GOU>

Query Match 28.6%; Score 2; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KP 5
 ||
 Db 4 KP 5

RESULT 21

S78764
ribosomal protein MRP-S23, mitochondrial - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: S78764
R;Graack, H.R.
submitted to the Protein Sequence Database, July 1999
A;Reference number: S78760
A;Accession: S78764
A;Molecule type: protein
A;Residues: 1-6 <GRA>
C;Keywords: mitochondrion
F;1-6/Product: ribosomal protein MRP-S23 (fragment) #status experimental <MAT>

Query Match 28.6%; Score 2; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KP 5
Db 5 KP 6

RESULT 22

A20186
fatty-acid synthase (EC 2.3.1.85) - rabbit (fragment)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 05-May-2000
C;Accession: A20186
R;McCarthy, A.D.; Atken, A.; Hardie, G.; Santikarn, S.; Williams, D.H.
FEBS Lett. 160, 296-300, 1983
A;Title: Amino acid sequence around the active serine in the acyl transferase domain of
A;Reference number: A20186; PMID:83287768; PMID:6554204
A;Accession: A20186
A;Molecule type: protein
A;Residues: 1-6 <MCC>
C;Keywords: acyltransferase; coenzyme A

Query Match 28.6%; Score 2; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GE 3
Db 3 GE 4

RESULT 23

B33932
Ig mu chain D region (D23) - mouse
C;Species: Mus musculus (house mouse)
C;Date: 09-Mar-1990 #sequence_revision 09-Mar-1990 #text_change 16-Aug-1996
C;Accession: B33932
R;Baccala, R.; Vo Quang, T.; Gilbert, M.; Terynck, T.; Avrameas, S.
Proc. Natl. Acad. Sci. U.S.A. 86, 4624-4628, 1989
A;Title: Two murine natural polyreactive autoantibodies are encoded by nonmutated germ-
A;Reference number: A33932; PMID:89282823; PMID:2499887
A;Accession: B33932
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-6 <BAC>
A;Cross-references: GB:M27107
C;Keywords: immunoglobulin

Query Match 28.6%; Score 2; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EK 4
Db 1 EK 2

RESULT 24

PT0518
T-cell receptor beta chain V-D-J region (100-4B) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0518
R;Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; PMID:91277601; PMID:1711558
A;Accession: PT0518
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-6 <FEE>
A;Experimental source: adult thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 28.6%; Score 2; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TG 2
Db 5 TG 6

RESULT 25

PT0630
T-cell receptor beta chain V-D-J region (111-1C) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0630
R;Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; PMID:91277601; PMID:1711558
A;Accession: PT0630
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-6 <FEE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 28.6%; Score 2; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TG 2
Db 5 TG 6

RESULT 26

PT0643
T-cell receptor beta chain V-D-J region (111-1E) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0643
R;Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; PMID:91277601; PMID:1711558
A;Accession: PT0643
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-6 <FEE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 28.6%; Score 2; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EK 4
Db 1 EK 2

QY 1 TG 2
||
Db 3 TG 4

RESULT 27

T-cell receptor beta chain V-D-J region (120-2CN) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0619; PT0563; PT0598
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0619
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-6 <FE3>
A;Experimental source: newborn thymus, strain BALB/c, clone 120-2CN
A;Accession: PT0563
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-6 <FE3>
A;Experimental source: day 18 fetal thymus, strain BALB/c, clone 136-3A
A;Accession: PT0598
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-6 <FE2>
A;Experimental source: newborn thymus, strain BALB/c, clone 111-1Q
C;Keywords: T-cell receptor

Query Match 28.6%; Score 2; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
||
Db 5 TG 6

RESULT 28

T-cell receptor beta chain V-D-J region (121-3BB) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0668
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0668
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-6 <FE3>
A;Experimental source: day 4 postnatal thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 28.6%; Score 2; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
||
Db 5 TG 6

RESULT 29

T-cell receptor beta chain V-D-J region (135-1AF) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997

C;Accession: PT0723

R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0723
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-6 <FE2>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 28.6%; Score 2; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
||
Db 4 TG 5

RESULT 30

T-cell receptor beta chain V-D-J region (161-2D) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0726
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0726
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-6 <FE2>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 28.6%; Score 2; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GE 3
||
Db 3 GE 4

RESULT 31

NYPG7
hypothalamic heptapeptide - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 01-Sep-1981 #sequence_revision 01-Sep-1981 #text_change 23-Aug-1996
C;Accession: A01417
R;Chang, R.C.C.; Huang, W.Y.; Arimura, A.; Redding, T.W.; Coy, D.H.; Saffran, M.; Kong
Horn. Metab. Res. 13, 228-232, 1981
A;Title: Isolation, structure and synthesis of a heptapeptide with in vitro ACTH-relea
A;Reference number: A01417; MUID:81213980; PMID:6263778
A;Accession: A01417
A;Molecule type: protein
A;Residues: 1-7 <CHA>
C;Superfamily: hypothalamic heptapeptide
C;Keywords: hypothalamus

Query Match 28.6%; Score 2; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 YK 7
||
Db 6 YK 7

RESULT 32

JN0859
 peptidyl-diesterase A inhibitory peptide C105 - striped bonito
 C:Species: Sarda orientalis (striped bonito)
 C:Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 07-May-1999
 C:Accession: JN0859
 R:Matsumura, N.; Fujii, M.; Takeda, Y.; Shimizu, T.
 Biosci. Biotechnol. Biochem. 57, 1743-1744, 1993
 A:Title: Isolation and characterization of angiotensin I-converting enzyme inhibitory peptide
 A:Reference number: JN0859; MUID:94080036; PMID:7764272
 A:Accession: JN0859
 A:Molecule type: protein
 A:Residues: 1-7 <MAT>
 A:Experimental source: intestine
 C:Comment: The carboxyl-terminus is essential for the protein's expression of angiotensin
 C:Superfamily: bradykinin-potentiating peptide
 C:Keywords: angiotensin-converting enzyme inhibitor

Query Match 28.6%; Score 2; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05; Mismatches 0; Indels 0; Gaps 0;
 Matches 2; Conservative 0;

QY 3 EK 4
 ||
 6 EX 7

Db

RESULT 33
 S16364
 opacity protein P.IIe - Neisseria gonorrhoeae (strain FA1090) (fragment)
 N:Alternate names: outer membrane protein P.IIe
 C:Species: Neisseria gonorrhoeae
 A:Variety: strain FA1090
 C:Date: 19-Mar-1997 #sequence_revision 17-Oct-1997 #text_change 17-Oct-1997
 C:Accession: S16364
 R:Barritt, D.S.; Schwalbe, R.S.; Klapper, D.G.; Cannon, J.G.
 Infect. Immun. 55, 2026-2031, 1987
 A:Title: Antigenic and structural differences among six proteins II expressed by a single
 A:Reference number: S16360; MUID:87306843; PMID:3114142
 A:Accession: S16364
 A:Molecule type: protein
 A:Residues: 1-7 <BAR>
 A:Experimental source: strain FA1090
 A:Note: expression of opacity proteins is regulated by the number of translated repeat
 of repeats place the start codon in frame with the rest of the protein
 C:Superfamily: opacity protein
 C:Keywords: cell surface component; transmembrane protein
 F:1-7/Product: opacity protein P.IIe (fragment) #status experimental <MAT>

Query Match 28.6%; Score 2; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05; Mismatches 0; Indels 0; Gaps 0;
 Matches 2; Conservative 0;

QY 2 GE 3
 ||
 2 GE 3

Db

RESULT 34
 A61081
 tryptophyllin, basic - Rohde's leaf frog
 C:Species: Phyllomedusa rohdei (Rohde's leaf frog)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 18-Aug-2000
 C:Accession: A61081
 R:Montecucchi, P.C.; Vincenti, M.; Lazzerini, A.M.; Rusconi, L.; Erspamer, V.
 Int. J. Pept. Protein Res. 33, 391-395, 1989
 A:Title: Isolation, structure determination and synthesis of a novel tryptophan-containing
 A:Reference number: A61081
 A:Accession: A61081
 A:Molecule type: protein
 A:Residues: 1-7 <MON>
 C:Comment: The biological activity of this peptide was not determined.
 C:Superfamily: unassigned animal peptides
 C:Keywords: amidated carboxyl end; hydroxyproline; skin

F:3/Modified site: 4-hydroxyproline (Pro) #status experimental
 F:7/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 28.6%; Score 2; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05; Mismatches 0; Indels 0; Gaps 0;
 Matches 2; Conservative 0;

QY 4 KP 5
 ||
 1 KP 2

Db

RESULT 35
 P00663
 membrane protein - porcine epidemic diarrhea virus (isolate Belgian CV777) (fragment)
 C:Species: porcine epidemic diarrhea virus
 C:Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 08-Oct-1999
 C:Accession: P00663
 R:Bridgen, A.; Duarte, M.; Tobler, K.; Laude, H.; Ackermann, M.
 J. Gen. Virol. 74, 1795-1804, 1993
 A:Title: Sequence determination of the nucleocapsid protein gene of the porcine epidemic
 issible gastroenteritis virus.
 A:Reference number: JQ2191; MUID:93389433; PMID:8397280
 A:Accession: P00663
 A:Molecule type: mRNA
 A:Residues: 1-7 <BRI>
 A:Cross-references: GB:Z14976; NID:g311650; PIDN:CAA78699.1; PID:g584083
 C:Comment: This virus is coronavirus related to human coronavirus 229E.
 C:Keywords: membrane protein

Query Match 28.6%; Score 2; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05; Mismatches 0; Indels 0; Gaps 0;
 Matches 2; Conservative 0;

QY 3 EK 4
 ||
 1 EK 2

Db

RESULT 36
 S45311
 microcin C7 - Escherichia coli plasmid pMccC7
 C:Species: Escherichia coli
 C:Date: 10-Dec-1994 #sequence_revision 24-May-1996 #text_change 17-Mar-1999
 C:Accession: S45311
 R:Gonzalez-Pastor, J.E.; San Millan, J.L.; Moreno, F.
 Nature 369, 281, 1994
 A:Title: The smallest known gene.
 A:Reference number: S45311; MUID:94239518; PMID:8183363
 A:Accession: S45311
 A:Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-7 <GON>
 C:Genetics:
 A:Gene: mcca
 A:Genome: plasmid pMccC7
 C:Keywords: antibacterial
 F:1/Modified site: N-formylmethionine #status predicted
 F:7/Modified site: asparagine derivative (Asn) #status experimental

Query Match 28.6%; Score 2; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05; Mismatches 0; Indels 0; Gaps 0;
 Matches 2; Conservative 0;

QY 1 TG 2
 ||
 3 TG 4

Db

RESULT 37
 A44428
 platelet aggregation-associated protein - Streptococcus sanguis (fragment)
 C:Species: Streptococcus sanguis

C>Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 11-Nov-1994
 C/Accession: A4428
 J. Brickson, P.R.; Herzberg, M.C.
 J. Biol. Chem. 268, 1646-1649, 1993
 A>Title: The Streptococcus sanguinis platelet aggregation-associated protein. Identification
 A/Reference number: A4428; MUID:93131902; PMID:8420939
 A/Accession: A4428
 A/Status: preliminary
 A/Molecule type: protein
 A/Residues: 1-7 <ERI>
 C/Keywords: cell wall

Query Match 28.6%; Score 2; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GE 3
 ||
 Db 2 GE 3

RESULT 38
 PS0254
 18K protein 5507 - rice (strain Nihonbare) (fragment)
 C/Species: Oryza sativa (rice)
 C/Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 23-Mar-1995
 C/Accession: PS0254
 R/Tsugita, A.
 submitted to JIPID, April 1993
 A/Reference number: PS0206
 A/Accession: PS0254
 A/Molecule type: protein
 A/Residues: 1-7 <TSU>
 A/Experimental source: leaf, chloroplast, strain Nihonbare
 A/Note: molecular weight 18K, pI 4.4

Query Match 28.6%; Score 2; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EK 4
 ||
 Db 6 EK 7

RESULT 39
 PT0611
 T-cell receptor beta chain V-D-J region (100-2AG) - mouse (fragment)
 C/Species: Mus musculus (house mouse)
 C/Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 C/Accession: PT0611
 R/Feeney, A.J.
 J. Exp. Med. 174, 115-124, 1991
 A>Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
 A/Reference number: PT0509; MUID:91277601; PMID:1711558
 A/Accession: PT0611
 A/Status: translation not shown
 A/Molecule type: mRNA
 A/Residues: 1-7 <FEE>
 A/Experimental source: newborn thymus, strain BALB/c
 C/Keywords: T-cell receptor

Query Match 28.6%; Score 2; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
 ||
 Db 6 TG 7

RESULT 40
 PT0526
 T-cell receptor beta chain V-D-J region (100-4E) - mouse (fragment)
 C/Species: Mus musculus (house mouse)
 C/Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 C/Accession: PT0526
 R/Feeney, A.J.
 J. Exp. Med. 174, 115-124, 1991
 A>Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
 A/Reference number: PT0509; MUID:91277601; PMID:1711558
 A/Accession: PT0526
 A/Status: translation not shown
 A/Molecule type: mRNA
 A/Residues: 1-7 <FEE>
 A/Experimental source: newborn thymus, strain BALB/c
 C/Keywords: T-cell receptor

Query Match 28.6%; Score 2; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EK 4
 ||
 Db 6 TG 7

T-cell receptor beta chain V-D-J region (100-4E) - mouse (fragment)
 C/Species: Mus musculus (house mouse)
 C/Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 C/Accession: PT0526
 R/Feeney, A.J.
 J. Exp. Med. 174, 115-124, 1991
 A>Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
 A/Reference number: PT0509; MUID:91277601; PMID:1711558
 A/Accession: PT0526
 A/Status: translation not shown
 A/Molecule type: mRNA
 A/Residues: 1-7 <FEE>
 A/Experimental source: adult thymus, strain BALB/c
 C/Keywords: T-cell receptor

Query Match 28.6%; Score 2; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GE 3
 ||
 Db 3 GE 4

RESULT 41
 PT0623
 T-cell receptor beta chain V-D-J region (111-1AB) - mouse (fragment)
 C/Species: Mus musculus (house mouse)
 C/Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 C/Accession: PT0623
 R/Feeney, A.J.
 J. Exp. Med. 174, 115-124, 1991
 A>Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
 A/Reference number: PT0509; MUID:91277601; PMID:1711558
 A/Accession: PT0623
 A/Status: translation not shown
 A/Molecule type: mRNA
 A/Residues: 1-7 <FEE>
 A/Experimental source: newborn thymus, strain BALB/c
 C/Keywords: T-cell receptor

Query Match 28.6%; Score 2; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
 ||
 Db 6 TG 7

RESULT 42
 PT0704
 T-cell receptor beta chain V-D-J region (135-1DU) - mouse (fragment)
 C/Species: Mus musculus (house mouse)
 C/Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 C/Accession: PT0704
 R/Feeney, A.J.
 J. Exp. Med. 174, 115-124, 1991
 A>Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
 A/Reference number: PT0509; MUID:91277601; PMID:1711558
 A/Accession: PT0704
 A/Status: translation not shown
 A/Molecule type: DNA
 A/Residues: 1-7 <FEE>
 A/Experimental source: newborn thymus, strain BALB/c
 C/Keywords: T-cell receptor

Query Match 28.6%; Score 2; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EK 4
 ||

```

Db          5 EK 6

RESULT 43
PT0689
T-cell receptor beta chain V-D-J region (140-1AF) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0689
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0689
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-7 <FEE>
A:Experimental source: day 18 fetal thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match      28.6%; Score 2; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TG 2
Db      6 TG 7

RESULT 44
PT0676
T-cell receptor beta chain V-D-J region (140-1AL) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0676
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0676
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-7 <FEE>
A:Experimental source: day 18 fetal thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match      28.6%; Score 2; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 GE 3
Db      3 GE 4

RESULT 45
PT0719
T-cell receptor beta chain V-D-J region (140-2F) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0719; PT0638
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0719
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-7 <FEE>
A:Experimental source: newborn thymus, strain BALB/c (clone 140-2F)
A:Accession: PT0638
A>Status: translation not shown
A:Molecule type: mRNA

A:Residues: 1-7 <FEE>
A:Experimental source: newborn thymus, strain BALB/c (clone 111-IN)
C:Keywords: T-cell receptor

Query Match      28.6%; Score 2; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TG 2
Db      5 TG 6

RESULT 46
PT0702
T-cell receptor beta chain V-D-J region (161-2AA) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0702
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0702
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-7 <FEE>
A:Experimental source: newborn thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match      28.6%; Score 2; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TG 2
Db      5 TG 6

RESULT 47
I55382
Hypothetical peptide PAII promoter region - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 16-Apr-1999 #sequence_revision 16-Apr-1999 #text_change 20-Apr-2000
C:Accession: I55382
R:Dawson, S.J.; Wiman, B.; Hamsten, A.; Green, F.; Humphries, S.; Henney, A.M.
J. Biol. Chem. 268, 10739-10745, 1993
A:Title: The two allele sequences of a common polymorphism in the promoter of the plasm
A:Reference number: I55382; MUID:9326509; PMID:838372
A:Accession: I55382
A>Status: translation not shown; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-7 <DAW>
A:Cross-references: GB:M91557; NID:gl90020; PIDD:AAA60110.1; PFD:gl90021
C:Comment: This is the hypothetical translation of a sequence from the PAII gene promot
C:Genetics:
A:Gene: GDB:PAII
A:Cross-references: GDB:120297; OMIM:173360
A:Map position: 7q21.3-7q22

Query Match      28.6%; Score 2; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 GE 3
Db      4 GE 5

RESULT 48
A28004
adipokinetic hormone G - two-spotted cricket
N:Alternate names: AKH-G

```

C:Species: Gryllus bimaculatus (two-spotted cricket)
C:Date: 30-Jun-1989 #sequence_revision 24-Oct-1997 #text_change 24-Oct-1997
C:Accession: A28004
R:Gaede, G.; Rinehart, K.L.
Biochem. Biophys. Res. Commun. 149, 908-914, 1987
A:Title: Primary sequence analysis by fast atom bombardment mass spectrometry of a peptide
A:Reference number: A28004; MUID:88106553; PMID:3426616
A:Accession: A28004
A:Molecule type: protein
A:Residues: 1-8 <GAE>
A>Note: the amino-terminal residue forms pyrrolidone carboxylic acid; therefore, we have
C:Superfamily: adipeptin-like hormone
C:Keywords: amidated carboxyl end; corpora cardiaca; hormone; neuropeptide; pyroglutamic
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:8/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match 28.6%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
||
Db 6 TG 7

RESULT 49
S20162
leghemoglobin III - Sesbania rostrata (fragment)
C:Species: Sesbania rostrata
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 04-Mar-2000
C:Accession: S20162
R:Metz, B.A.; Welters, P.; Hoffmann, H.J.; Jensen, E.O.; Schell, J.; de Bruijn, F.J.
Mol. Gen. Genet. 214, 181-191, 1988
A:Title: Primary structure and promoter analysis of leghemoglobin genes of the stem-nodu
A:Reference number: S08322; MUID:89181515; PMID:3237206
A:Accession: S20162
A:Molecule type: DNA
A:Residues: 1-8 <MET>
A:Cross-references: EMBL:X13504; NID:g21383; PIDN:CAA31850.1; PID:g579482
C:Genetics:
C:Gene: glb3
C:Superfamily: globin; globin homology
C:Keywords: heme; oxygen carrier

Query Match 28.6%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EK 4
||
Db 5 EK 6

RESULT 50
S37141
rpsA protein - Erwinia chrysanthemi
C:Species: Erwinia chrysanthemi
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Oct-1999
C:Accession: S37141
R:Douillie, A.; Toussaint, A.; Faelen, M.
submitted to the EMBL Data Library, August 1993
A:Description: Identification of the integration host factor genes of E. chrysanthemi.
A:Reference number: S37139
A:Accession: S37141
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-8 <DOU>
A:Cross-references: EMBL:X74750; NID:g399669; PIDN:CAA52769.1; PID:g581108

Query Match 28.6%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

C:Species: Gryllus bimaculatus (two-spotted cricket)
C:Date: 30-Jun-1989 #sequence_revision 24-Oct-1997 #text_change 24-Oct-1997
C:Accession: A28004
R:Gaede, G.; Rinehart, K.L.
Biochem. Biophys. Res. Commun. 149, 908-914, 1987
A:Title: Primary sequence analysis by fast atom bombardment mass spectrometry of a peptide
A:Reference number: A28004; MUID:88106553; PMID:3426616
A:Accession: A28004
A:Molecule type: protein
A:Residues: 1-8 <GAE>
A>Note: the amino-terminal residue forms pyrrolidone carboxylic acid; therefore, we have
C:Superfamily: adipeptin-like hormone
C:Keywords: amidated carboxyl end; corpora cardiaca; hormone; neuropeptide; pyroglutamic
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:8/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match 28.6%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
||
Db 6 TG 7

RESULT 51
S63493
disimilatory sulfite reductase gamma chain, membrane-bound and soluble - Desulfovibri
C:Species: Desulfovibrio desulfuricans
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
C:Accession: S63493; S63494
R:Steuber, J.; Arendsen, A.F.; Hagen, W.R.; Kroneck, P.M.H.
Eur. J. Biochem. 233, 873-879, 1995
A:Title: Molecular properties of the dissimilatory sulfite reductase from Desulfovibri
A:Reference number: S63489; MUID:96085152; PMID:8521853
A:Accession: S63493
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-8 <STE>
A:Accession: S63494
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-8 <ST2>

Query Match 28.6%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 YK 7
||
Db 5 YK 6

RESULT 52
A39308
glycine reductase (EC 1.4.99.-) sulfhydryl protein C, alpha chain - Clostridium stickl
C:Species: Clostridium sticklandii
C:Date: 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change 15-Aug-1997
C:Accession: A39308
R:Stadtman, T.C.; Davis, J.N.
J. Biol. Chem. 266, 22147-22153, 1991
A:Title: Glycine reductase protein C. Properties and characterization of its role in t
A:Reference number: A39308; MUID:92042141; PMID:1939235
A:Accession: A39308
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-8 <STA>
C:Function:
A:Description: glycine reductase complex catalyzes the reductive deamination of glycin
C:Keywords: ATP; oxidoreductase

Query Match 28.6%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KP 5
||
Db 2 KP 3

RESULT 53
S68802
nitrate reductase (NADH) inhibitor - spinach (fragment)
C:Species: Spinacia oleracea (spinach)
C:Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 30-Jan-1998
C:Accession: S68802
R:Bachmann, N.; Huber, J.L.; Liao, P.C.; Gage, D.A.; Huber, S.C.
FEBS Lett. 387, 127-131, 1996
A:Title: The inhibitor protein of phosphorylated nitrate reductase from spinach (spina
A:Reference number: S68802; MUID:96244508; PMID:8674533
A:Accession: S68802
A:Molecule type: protein
A:Residues: 1-8 <BAC>

A;Experimental source: leaves; strain cv. Bloomsdale

Query Match 28.6%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 YK 7
||
Db 7 YK 8

RESULT 54

S16324
hypothetical protein 2 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 21-Nov-1993 #sequence_revision 12-May-1995 #text_change 21-Jul-2000
C;Accession: S16324
R;Ruberti, I.; Sessa, G.; Lucchetti, S.; Morelli, G.
EMBO J. 10, 1787-1791, 1991

A;Title: A novel class of plant proteins containing a homeodomain with a closely linked
A;Reference number: S16323; MUID:91266907; PMID:1675603
A;Accession: S16324

A;Status: translation not shown

A;Molecule type: mRNA

A;Residues: 1-8 <RUB>

A;Cross-References: EMBL:X58821; NID:gl6327; PIDN:CAA41624.1; PID:g579259

Query Match 28.6%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 YK 7
||
Db 3 YK 4

RESULT 55

PQ0701

unidentified 6 S/31K protein [imported] - rice (fragment)

C;Species: Oryza sativa (rice)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C;Accession: PQ0701

R;Komatsu, S.; Kajiwara, H.; Hirano, H.

Theor. Appl. Genet. 86, 935-942, 1993

A;Title: A rice protein library; a data-file of rice proteins separated by two-dimension

A;Reference number: PQ0696

A;Accession: PQ0701

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-8 <KOW>

Query Match 28.6%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
||
Db 6 TG 7

RESULT 56

PT0030

inulinase (EC 3.2.1.7) - Aspergillus ficuum (fragment)

N;Alternate names: inulase

C;Species: Aspergillus ficuum

C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 07-May-1999

C;Accession: PT0030

R;Ettaib, M.; Baratti, J.C.

Agric. Biol. Chem. 54, 61-68, 1990

A;Title: Molecular and kinetic properties of Aspergillus ficuum inulinases.

A;Reference number: PT0030; MUID:90344234; PMID:1368526

A;Accession: PT0030

A;Molecule type: protein

A;Residues: 1-8 <ETT>

C;Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 28.6%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PY 6
||
Db 6 PY 7

RESULT 57

D61512

variant surface glycoprotein M1rat 1.6 - Trypanosoma brucei (fragment)

C;Species: Trypanosoma brucei

C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 07-May-1999

C;Accession: D61512

R;Holder, A.A.; Cross, G.A.M.

Mol. Biochem. Parasitol. 2, 135-150, 1981

A;Title: Glycopeptides from variant surface glycoproteins of Trypanosoma brucei. C-term
A;Reference number: A61512; MUID:81172836; PMID:6163983

A;Accession: D61512

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-8 <HOL>

C;Keywords: glycoprotein

Query Match 28.6%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GE 3
||
Db 3 GE 4

RESULT 58

D47393

neuropeptide callatostatin 4 - bluebottle fly (Calliphora vomitoria)

C;Species: Calliphora vomitoria

C;Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 28-Apr-1995

C;Accession: D47393

R;Duve, H.; Johnsen, A.H.; Scott, A.G.; Yu, C.G.; Yagi, K.J.; Tobe, S.S.; Thorpe, A.

Proc. Natl. Acad. Sci. U.S.A. 90, 2456-2460, 1993

A;Title: Callatostatins: neuropeptides from the blowfly Calliphora vomitoria with sequ

A;Reference number: A47393; MUID:93211980; PMID:8460157

A;Accession: D47393

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-8 <DUV>

A;Experimental source: thoracic ganglia

A;Note: sequence extracted from NCBI backbone (NCBIP:128479)

Query Match 28.6%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PY 6
||
Db 3 PY 4

RESULT 59

E47393

neuropeptide callatostatin 5 - bluebottle fly (Calliphora vomitoria)

C;Species: Calliphora vomitoria

C;Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 03-Mar-1995

C;Accession: E47393

R;Duve, H.; Johnsen, A.H.; Scott, A.G.; Yu, C.G.; Yagi, K.J.; Tobe, S.S.; Thorpe, A.

Proc. Natl. Acad. Sci. U.S.A. 90, 2456-2460, 1993

A;Title: Callatostatins: neuropeptides from the blowfly Calliphora vomitoria with sequ

A;Reference number: A47393; MUID:93211980; PMID:8460157

A:Accession: E47393
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-8 <DUV>
A:Experimental source: whole flies
A>Note: sequence extracted from NCBI backbone (NCBIP:128482)

Query Match 28.6%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PY 6
||
3 PY 4

Db

RESULT 60
B45800
serum albumin - dog (fragment)
C:Species: Canis lupus familiaris (dog)
C:Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 31-Dec-1993
C:Accession: B45800
R:Caraway, R.E.; Cochrane, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A:Title: Structures of histamine-releasing peptides formed by the action of acid proteases
A:Reference number: A45800; MUID:89341406; PMID:2474609
A:Accession: B45800
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-8 <CAR>

Query Match 28.6%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PY 6
||
6 PY 7

Db

RESULT 61
I48934
apolipoprotein A-II - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C:Accession: I48934
R:Ko, M.S.; Wang, X.; Horton, J.H.; Hagen, M.D.; Takahashi, N.; Maerzaki, Y.; Nadeau, J.H.
Mamm. Genome 5, 349-355, 1994
A:Title: Genetic mapping of 40 cDNA clones on the mouse genome by PCR.
A:Reference number: I48934; MUID:94319082; PMID:8043949
A:Accession: I48934
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-8 <RES>
A:Cross-references: EMBL:U05691; NID:g497010; PIDN:AAB60462.1; PID:g642826

Query Match 28.6%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GE 3
||
1 GE 2

Db

RESULT 62
A54823
olfactory receptor 17 - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 17-Mar-1999
C:Accession: A54823
R:Chess, A.; Simon, I.; Cedar, H.; Axel, R.
Cell 78, 823-834, 1994

A:Title: Allelic inactivation regulates olfactory receptor gene expression.
A:Reference number: A54823; MUID:94373818; PMID:8087849
A:Accession: A54823
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-8 <CHE>

Query Match 28.6%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
||
7 TG 8

Db

RESULT 63
PT0595
T-cell receptor beta chain V-D-J region (100-2AA) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0595
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0595
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-8 <FEE>
A:Experimental source: newborn thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match 28.6%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
||
6 TG 7

Db

RESULT 64
PT0530
T-cell receptor beta chain V-D-J region (100-4AK) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0530
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0530
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-8 <FEE>
A:Experimental source: adult thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match 28.6%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
||
6 TG 7

Db

RESULT 65
PT0639
T-cell receptor beta chain V-D-J region (111-1AA) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997

C:Accession: PT0639

R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A:Reference number: PT0509; MUID:91277601; PMID:1711558

C:Accession: PT0639

A:Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-8 <PEE>

A:Experimental source: newborn thymus, strain BALB/c

C:Keywords: T-cell receptor

Query Match 28.6%; Score 2; DB 2; Length 8;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2

Db 4 TG 5

RESULT 66

PT0631

T-cell receptor beta chain V-D-J region (111-11) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997

C:Accession: PT0631

R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A:Reference number: PT0509; MUID:91277601; PMID:1711558

C:Accession: PT0631

A:Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-8 <PEE>

A:Experimental source: newborn thymus, strain BALB/c

C:Keywords: T-cell receptor

Query Match 28.6%; Score 2; DB 2; Length 8;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2

Db 7 TG 8

RESULT 67

PT0653

T-cell receptor beta chain V-D-J region (121-3H) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997

C:Accession: PT0653

R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A:Reference number: PT0509; MUID:91277601; PMID:1711558

C:Accession: PT0653

A:Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-8 <PEE>

A:Experimental source: day 4 postnatal thymus, strain BALB/c

C:Keywords: T-cell receptor

Query Match 28.6%; Score 2; DB 2; Length 8;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GE 3

Db 6 GE 7

RESULT 68

PT0588

T-cell receptor beta chain V-D-J region (141-LCB) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997

C:Accession: PT0588

R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A:Reference number: PT0509; MUID:91277601; PMID:1711558

C:Accession: PT0588

A:Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-8 <PEE>

A:Experimental source: day 19 fetal thymus, strain BALB/c

C:Keywords: T-cell receptor

Query Match 28.6%; Score 2; DB 2; Length 8;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2

Db 6 TG 7

RESULT 69

PT0691

T-cell receptor beta chain V-D-J region (154-2K) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997

C:Accession: PT0691

R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A:Reference number: PT0509; MUID:91277601; PMID:1711558

C:Accession: PT0691

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-8 <PEE>

A:Experimental source: day 18 fetal thymus, strain BALB/c

C:Keywords: T-cell receptor

Query Match 28.6%; Score 2; DB 2; Length 8;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2

Db 7 TG 8

RESULT 70

C39690

neural cell adhesion molecule, cardiac splice form -, -, -, - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C:Date: 24-Jan-1992 #sequence_revision 24-Jan-1992 #text_change 16-Jul-1999

C:Accession: C39690

R:Reyes, A.A.; Small, S.J.; Axelson, R.

Mol. Cell. Biol. 11, 1654-1661, 1991

A:Title: At least 27 alternatively spliced forms of the neural cell adhesion molecule m

A:Reference number: A39690; MUID:9141516; PMID:1996115

A:Accession: C39690

A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tr

A:Molecule type: mRNA

A:Residues: 1-8 <REV>

A:Cross-references: GB:M63570

C:Keywords: cardiac muscle; cell adhesion; heart

Query Match 28.6%; Score 2; DB 2; Length 8;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GE 3
||
5 GE 6

RESULT 71

B47594
aspartate kinase (EC 2.7.2.4) beta chain - Corynebacterium flavum (strain N13) (fragment)
C:Species: Corynebacterium flavum
C:Date: 19-Mar-1997 #sequence_revision 19-Dec-1997 #text_change 19-Dec-1997
C:Accession: B47594
R:Pollettie, M.T.; Peoples, O.P.; Agotopoulou, C.; Sinskey, A.J.
J. Bacteriol. 175, 4056-4103, 1993
A:Title: Gene structure and expression of the Corynebacterium flavum N13 ask-ase operon.
A:Reference number: A47594; MUID:93308089; PMID:8100567
A:Accession: B47594
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-8 <FOL>
C:Keywords: phosphotransferase

Query Match 28.6%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
||
7 TG 8

RESULT 72

S07241
litorin - Rohde's leaf frog
C:Species: Phyllomedusa rohdei (Rohde's leaf frog)
C:Date: 12-Feb-1993 #sequence_revision 12-Mar-1993 #text_change 18-Aug-2000
C:Accession: S07241
R:Barra, D.; Falconieri Erspamer, G.; Simmaco, M.; Bossa, F.; Melchiorri, P.; Erspamer, F.E.S. Lett. 182, 53-56, 1985
A:Title: Rohdei-litorin: a new peptide from the skin of Phyllomedusa rohdei.
A:Reference number: S07241; MUID:85127560; PMID:3838283
A:Accession: S07241
A:Molecule type: protein
A:Residues: 1-9 <BAR>
C:Superfamily: gastrin-releasing peptide
C:Keywords: amidated carboxyl end; blocked amino end; neuropeptide; pyroglutamic acid

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:9/Modified site: amidated carboxyl end (Met) #status experimental

Query Match 28.6%; Score 2; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
||
5 TG 6

RESULT 73

A61357
phyllocaerulein - Sauvage's leaf frog
C:Species: Phyllomedusa sauvagei (Sauvage's leaf frog)
C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 02-Sep-2000
C:Accession: A61357
R:Anastasi, A.; Bertaccini, G.; Cei, J.M.; De Caro, G.; Erspamer, V.; Impicciatore, M. Br. J. Pharmacol. 37, 198-206, 1969
A:Title: Structure and pharmacological actions of phyllocaerulein, a caerulein-like nona
A:Reference number: A61357; MUID:70005484; PMID:5824931
A:Accession: A61357
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-9 <ANA>
C:Superfamily: gastrin

C:Keywords: amidated carboxyl end; neuropeptide; pyroglutamic acid; skin; sulfoprotein

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:3/Binding site: sulfate (Tyr) (covalent) #status experimental
F:9/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 28.6%; Score 2; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
||
4 TG 5

RESULT 74

S13636
coat protein beta chain, Golgi-derived - rabbit (fragment)
N:Alternate names: beta-COP protein
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 19-Mar-1997 #sequence_revision 24-Mar-1999 #text_change 24-Mar-1999
C:Accession: S13636
R:Serafini, T.; Stenbeck, G.; Brecht, A.; Lottspeich, F.; Orci, L.; Rothman, J.E.; Wie
Nature 349, 215-220, 1991
A:Title: A coat subunit of Golgi-derived non-clathrin-coated vesicles with homology to
A:Reference number: S13636; MUID:91101693; PMID:1898984
A:Accession: S13636
A:Molecule type: protein
A:Residues: 1-9 <SER>
C:Superfamily: coatamer complex beta chain
C:Keywords: Golgi apparatus; protein transport

Query Match 28.6%; Score 2; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GE 3
||
3 GE 4

RESULT 75

QDRB
delta sleep-inducing peptide - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 20-Jun-2000 #sequence_revision 20-Jun-2000 #text_change 20-Jun-2000
C:Accession: A01422
R:Monnier, M.; Dudler, L.; Gachter, R.; Maier, P.F.; Tobler, H.J.; Schoenenberger, G.A
Experientia 33, 548-552, 1977
A:Title: The delta sleep inducing peptide (DSIP). Comparative properties of the origin
A:Reference number: A01422; MUID:77185324; PMID:862769
A:Accession: A01422
A:Molecule type: protein
A:Residues: 1-9 <MON>
C:Comment: This peptide was obtained from dialysates of occipital venous sinus blood f;
of recipient rabbits, it induces spindle and delta EEG activity and reduced motor act;
C:Superfamily: unassigned animal peptides

Query Match 28.6%; Score 2; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GE 3
||
8 GE 9

RESULT 76

S66608
quinoline 2-oxidoeductase gamma chain - Comamonas testosteroni (fragment)
C:Species: Comamonas testosteroni
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
C:Accession: S66608
R:Schach, S.; Tshisuaka, B.; Fetzner, S.; Lingens, F.
Eur. J. Biochem. 232, 536-544, 1995

A;Title: Quinolone 2-oxidoreductase and 2-oxo-1,2-dihydroquinoline 5,6-dioxygenase from
 A;Reference number: S66606; MUID:96035889; PMID:7556204
 A;Accession: S66608
 A;Molecule type: protein
 A;Residues: 1-9 <SCH>
 A;Experimental source: strain 63

Query Match 28.6%; Score 2; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EK 4
 ||
 Db 5 EX 6

RESULT 77

S63491
 dissimilatory sulfite reductase beta chain, membrane-bound - Desulfovibrio desulfuricans
 C;Species: Desulfovibrio desulfuricans
 C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
 C;Accession: S63491

R;Steuber, J.; Arendsen, A.F.; Hagen, W.R.; Kroneck, P.M.H.
 Eur. J. Biochem. 233, 873-879, 1995

A;Title: Molecular properties of the dissimilatory sulfite reductase from Desulfovibrio

A;Reference number: S63489; MUID:96085152; PMID:8521853

A;Accession: S63491

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-9 <STE>

Query Match 28.6%; Score 2; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
 ||
 Db 5 TG 6

RESULT 78

A43848
 cell surface adhesin for heparan sulfate, 66K - Staphylococcus aureus (fragment)
 C;Species: Staphylococcus aureus

C;Date: 10-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 24-Feb-1995

C;Accession: A43848

R;Liang, O.D.; Ascencio, F.; Fransson, L.A.; Wadstrom, T.

Infect. Immun. 60, 899-906, 1992

A;Title: Binding of heparan sulfate to Staphylococcus aureus.

A;Reference number: A43848; MUID:92176005; PMID:1541563

A;Accession: A43848

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-9 <LIA>

A;Note: sequence extracted from NCBI backbone (NCBIP:85442)

Query Match 28.6%; Score 2; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
 ||
 Db 3 TG 4

RESULT 79

S66419
 tetrameric proteinase precursor, 60K, dithiothreitol-sensitive - spinach (fragment)
 C;Species: Spinacia oleracea (spinach)

C;Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 29-Aug-1997

C;Accession: S66419

R;Kuwabara, T.

FEBS Lett. 371, 195-198, 1995

A;Title: The 60-kDa precursor to the dithiothreitol-sensitive tetrameric protease of sp
 A;Reference number: S66419; MUID:95402209; PMID:7672127
 A;Accession: S66419
 A;Molecule type: protein
 A;Residues: 1-9 <KUN>

Query Match 28.6%; Score 2; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EK 4
 ||
 Db 8 EK 9

RESULT 80

PA0033

protein QA300033 - Arabidopsis thaliana (fragment)

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 30-Jun-1992 #sequence_revision 06-Jan-1995 #text_change 06-Jun-1997

C;Accession: PA0033

R;Kamo, M.; Kawakami, T.; Miyatake, N.; Tsugita, A.

submitted to JIPID, July 1994

A;Description: Separation and characterization of Arabidopsis proteins by two-dimension.

A;Reference number: PA0001

A;Accession: PA0033

A;Molecule type: protein

A;Residues: 1-9 <KAW>

A;Experimental source: leaf

Query Match 28.6%; Score 2; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GE 3
 ||
 Db 4 GE 5

RESULT 81

S70332

endosperm protein, 10K - rye (fragment)

C;Species: Secale cereale (rye)

C;Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 17-Apr-1998

C;Accession: S70332

R;Rocher, A.; Calero, M.; Soriano, F.; Mendez, E.

Biochim. Biophys. Acta 1295, 13-22, 1996

A;Title: Identification of major rye scalins as coeliac immunoreactive proteins.

A;Reference number: S70327; MUID:96283789; PMID:8679669

A;Accession: S70332

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-9 <ROC>

Query Match 28.6%; Score 2; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
 ||
 Db 1 TG 2

RESULT 82

T31612

hypothetical protein Y50E8A.h - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C;Accession: T31612

R;Steward, C.

submitted to the EMBL Data Library, September 1999

A;Reference number: Z21047

A;Accession: T31612

```
C:\date: 17-Sep-1997 #sequence_#revision 17-Sep-1997 #text_change 30-Jan-1998
```

C;Accession: S77984
 R;Arnold, S.; Lee, J.; Kim, M.; Song, E.; Linder, D.; Lottspeich, F.; Kadenbach, B.
 Submitted to the Protein Sequence Database, June 1997
 A;Reference number: S77980
 A;Accession: S77984
 A;Molecule type: protein
 A;Residues: 1-9 <ARN>
 A;Experimental source: heart
 C;Genetics:
 A;Genome: nuclear
 C;Function:
 A;Pathway: oxidative phosphorylation; respiratory chain
 C;Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane

Query Match 28.6%; Score 2; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PY 6
 ||
 DB 8 PY 9

RESULT 89

A53797
 C;Species: Rana catesbeiana (bullfrog)
 C;Date: 01-Dec-1995 #sequence_revision 01-Dec-1995 #text_change 17-Mar-1999
 C;Accession: A53797
 R;Tsuboi, S.; Matsumoto, H.; Jackson, K.W.; Tsujimoto, K.; Williams, T.; Yamazaki, A.
 J. Biol. Chem. 269, 15016-15023, 1994
 A;Title: Phosphorylation of an inhibitory subunit of cGMP phosphodiesterase in Rana catesbeiana
 A;Reference number: A53797; PMID:94253058; PMID:8195138

A;Accession: A53797
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 1-9 <TSU>
 C;Keywords: cGMP binding; phosphoric diester hydrolase

Query Match 28.6%; Score 2; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
 ||
 DB 2 TG 3

RESULT 90

JQ0914
 MHC class I histocompatibility antigen heavy chain - chicken (fragment)
 C;Species: Gallus gallus (chicken)
 C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 24-Feb-1995
 C;Accession: JQ0914
 R;Kanki, T.; Kuwasawa, N.; Sekiya, Y.; Ichikawa, Y.
 submitted to JIPIID, May 1991
 A;Description: Responsive expression of a MHC class I epitope and genes following Marek's disease virus infection
 A;Reference number: JQ0914
 A;Accession: JQ0914
 A;Molecule type: mRNA
 A;Residues: 1-9 <KAN>
 A;Experimental source: kidney, strain cornell N

Query Match 28.6%; Score 2; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TG 2
 ||
 DB 3 TG 4

RESULT 91

A60356
 I18K stomach cancer antigen - human (fragment)
 C;Species: Homo sapiens (man)
 C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 17-Mar-1999
 C;Accession: A60356
 R;Shiraishi, Y.
 Int. J. Cancer 45, 783-787, 1990
 A;Title: Western blotting analysis for malignant lymphoma and stomach cancer antigens f
 A;Reference number: A60356; MUID:90216080; PMID:2323853

A;Accession: A60356
 A;Molecule type: protein
 A;Residues: 1-9 <SHI>
 C;Keywords: glycoprotein

Query Match 28.6%; Score 2; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KP 5
 ||
 DB 4 KP 5

RESULT 92

PT0315
 Ig heavy chain CRD3 region (clone 6-109) - human (fragment)
 C;Species: Homo sapiens (man)
 C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
 C;Accession: PT0315
 R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
 J. Exp. Med. 173, 395-407, 1991
 A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and
 A;Reference number: PT0222; MUID:91108337; PMID:1899102

A;Accession: PT0315
 A;Molecule type: DNA
 A;Residues: 1-9 <YAM>
 A;Experimental source: B lymphocyte
 C;Keywords: heterotetramer; immunoglobulin

Query Match 28.6%; Score 2; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GE 3
 ||
 DB 1 GE 2

RESULT 93

PT0324
 Ig heavy chain CRD3 region (clone J2-106C) - human (fragment)
 C;Species: Homo sapiens (man)
 C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
 C;Accession: PT0324
 R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
 J. Exp. Med. 173, 395-407, 1991
 A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and
 A;Reference number: PT0222; MUID:91108337; PMID:1899102

A;Accession: PT0324
 A;Molecule type: DNA
 A;Residues: 1-9 <YAM>
 A;Experimental source: B lymphocyte
 C;Keywords: heterotetramer; immunoglobulin

Query Match 28.6%; Score 2; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GE 3
 ||
 DB 5 GE 6

RESULT 94
A61386
macrophage inhibitory factor (F5 cells) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 07-May-1999
C:Accession: A61386
R:Okii, S.; Hirose, S.I.; Higuchi, M.; Osawa, T.
Lymphokine Cytokine Res. 10, 273-280, 1991
A:Title: Macrophage migration inhibitory factor (MIF) produced by a human T cell hybrid
A:Reference number: A61386; MUID:92032107; PMID:1932371
A:Accession: A61386
A:Status: Preliminary
A:Molecule type: protein
A:Residues: 1-9 <OKI>

Query Match 28.6%; Score 2; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EK 4
||
DB 8 EK 9

RESULT 95
PC2197
zymogen granule membrane associated protein, ZAP36 - dog (fragment)
C:Species: Canis lupus familiaris (dog)
C:Date: 29-Aug-1994 #sequence_revision 05-Apr-1995 #text_change 07-May-1999
C:Accession: PC2197
R:Fukuoka, S.
Biosci. Biotechnol. Biochem. 58, 1282-1285, 1994
A:Title: Analysis of ZAPs, zymogen granule membrane associated proteins, in the regulated secretion of canine zymogen granules
A:Reference number: PC2197; MUID:94362286; PMID:7765250
A:Accession: PC2197
A:Molecule type: protein
A:Residues: 1-9 <FUK>
A:Experimental source: pancreas
A:Comment: This protein localizes in the zymogen granule membrane and has functions at the apical surface of the zymogen granule membrane
C:Keywords: membrane protein

Query Match 28.6%; Score 2; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KP 5
||
DB 8 KP 9

RESULT 96
I46016
cytokerin 4 - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 05-Nov-1999
C:Accession: I46016
R:Blessing, M.; Jorcano, J.L.; Franke, W.W.
EMBO J. 8, 117-126, 1989
A:Title: Enhancer elements directing cell-type-specific expression of cytokeratin genes
A:Reference number: I46016; MUID:89231609; PMID:2469572
A:Accession: I46016
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-9 <BLE>
A:Cross-references: EMBL:X14478; NID:g303; PIDN:CAA32640.1; PID:g577897

Query Match 28.6%; Score 2; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 YK 7
||
DB 3 YK 4

RESULT 97
S10784
enamelin i - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 19-Mar-1997 #sequence_revision 21-Nov-1998 #text_change 21-Nov-1998
C:Accession: S10784
R:Strawich, E.; Glimcher, M.J.
Eur. J. Biochem. 191, 47-56, 1990
A:Title: Tooth 'enamelins' identified mainly as serum proteins. Major 'enamelin' is all alpha-enamelin
A:Reference number: S10780; MUID:90336641; PMID:2379503
A:Accession: S10784
A:Molecule type: protein
A:Residues: 1-9 <STR>
C:Keywords: enamel; phosphoprotein

Query Match 28.6%; Score 2; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PY 6
||
DB 2 PY 3

RESULT 98
PH1591
Ig H chain V-D-J region (wild-type clone 142) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999
C:Accession: PH1591
R:Levinson, D.A.; Campos-Torres, J.; Leder, P.
J. Exp. Med. 178, 317-329, 1993
A:Title: Molecular characterization of transgene-induced immunodeficiency in B-less mice
A:Reference number: PH1580; MUID:93301609; PMID:8315387
A:Accession: PH1591
A:Molecule type: DNA
A:Residues: 1-9 <LEV>
A:Experimental source: bone marrow pre-B lymphocyte
C:Keywords: immunoglobulin

Query Match 28.6%; Score 2; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PY 6
||
DB 6 PY 7

RESULT 99
PH0108
late G1-69 protein - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C:Accession: PH0108
R:Nikaido, T.; Bradley, D.W.; Pardee, A.B.
Exp. Cell Res. 192, 102-109, 1991
A:Title: Molecular cloning of transcripts that accumulate during the late G1 phase in c-myc overexpressing cells
A:Reference number: PH0108; MUID:91078351; PMID:1984406
A:Accession: PH0108
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-9 <NIK>

Query Match 28.6%; Score 2; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CE 3
||
DB 6 CE 7

RESULT 100
I49406
bone gla protein - western wild mouse (fragment)
C:Species: Mus spretus (western wild mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C:Accession: I49406
R:Ko, M.S.; Wang, X.; Horton, J.H.; Hagen, M.D.; Takahashi, N.; Maezaki, Y.; Nadeau, J.H.
Mamm. Genome 5, 349-355, 1994
A:Title: Genetic mapping of 40 cDNA clones on the mouse genome by PCR.
A:Reference number: I48934; MUID:94319082; PMID:8043949
A:Accession: I49406
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-9 <RES>
A:Cross-references: EMBL:U05695; NID:G497015; PIDN:AAB60465.1; PID:G642828

Query Match 28.6%; Score 2; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 6 YK 7
Db 1 YK 2

Search completed: October 21, 2003, 18:29:06
Job time : 26.5556 sec

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 17:47:49 ; Search time 22.922 Seconds
(without alignments)
578.975 Million cell updates/sec

Title: US-10-057-890A-10

Perfect score: 797

Sequence: 1 MDYQVSSPIYDINYITSEPC.....GLNCCSSNRLLDGHQHVAA 138

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 76:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	379	47.6	352	2 A43113	chemokine (C-C) re
2	154	19.3	201	2 I57505	zinc finger protei
3	144	18.1	693	2 I37570	zinc finger protei
4	143.5	18.0	319	2 T46469	hypothetical prote
5	141	17.7	555	2 I53869	zinc finger protei
6	138.5	17.4	223	2 P31201	GLI-related finger
7	136.5	17.1	348	2 I38599	zinc finger protei
8	135.5	17.0	572	2 I39311	kruppel-type zinc
9	134	16.8	710	2 I48668	zinc finger protei
10	133.5	16.8	475	2 S03679	finger protein (cl
11	133.5	16.8	636	2 I48689	gene NK10 protein
12	133.5	16.8	701	2 T14757	hypothetical prote
13	131.5	16.5	686	2 I34612	zinc finger protei
14	131	16.4	261	2 S70006	finger protein zfo
15	129.5	16.2	614	2 JH0500	zinc finger protei
16	128.5	16.1	386	2 T12527	hypothetical prote
17	128.5	16.1	580	2 A37107	spermatogenesis pr
18	127.5	16.0	469	2 I38600	zinc finger protei
19	127.5	16.0	1173	2 I50620	procr2 - chicken
20	127	15.9	728	2 A48830	probable transcrip
21	126.5	15.9	732	2 S47073	finger protein HZF
22	124	15.6	367	2 S06582	finger protein (cl
23	124	15.6	599	2 I38598	zinc finger protei
24	123.5	15.5	209	2 S47068	finger protein HZF
25	123.5	15.5	347	2 S00549	developmental cont
26	123.5	15.5	378	2 S33994	finger protein ZNF
27	122.5	15.4	488	2 S47072	finger protein HZF
28	122.5	15.4	594	2 T12488	hypothetical prote
29	122.5	15.4	803	2 S26823	zinc finger protei

30 122 15.3 1350 2 S00647
31 120.5 15.1 189 2 A39240
32 120.5 15.1 229 2 A48927
33 120.5 15.1 399 2 S47071
34 120.5 15.1 546 2 I49636
35 120.5 15.1 1191 2 S35305
36 120 15.1 435 2 S00833
37 120 15.1 654 2 A57785
38 119.5 15.0 194 2 I53859
39 119.5 15.0 337 2 S06520
40 119.5 15.0 393 2 JN0533
41 119.5 15.0 576 2 A48157
42 119.5 15.0 671 2 JE0288
43 118.5 14.9 107 2 I39315
44 118.5 14.9 120 2 G02493
45 118.5 14.9 183 2 S70007

finger protein - A
finger protein mfg
kruppel-like zinc
finger protein HZF
DNA-binding protei
zinc finger protei
finger protein (cl
finger protein ZNF
zinc finger protei
finger protein PML
renal transcriptio
kruppel-type zinc
zinc-finger protei
KR-ZNF1 - human (f
finger protein zfo

ALIGNMENTS

RESULT 1

A43113

Chemokine (C-C) receptor 5 - human

N:Alternate names: C-C CKR-5; CCR5

C:Species: Homo sapiens (man)

C:Date: 12-Jul-1996 #sequence revision 12-Jul-1996 #text change 20-Jun-2000

C:Accession: A43113; S71808; A58834; A58832; G02653; A58833

R:Samson, M.; Labbe, O.; Mollereau, C.; Vassart, G.; Parmentier, M.

Biochemistry 35, 3362-3367, 1996

A:Title: Molecular cloning and functional expression of a new human CC-chemokine recei

A:Reference number: A43113; MUID:96241590; PMID:8639485

A:Accession: A43113

A:Molecule type: mRNA

A:Residues: 1-352 <SAM1>

A:Cross-references: GB:X91492; NID:gl262810; PIDN:CAA62796.1; PID:gl262811

R:Samson, M.; Libert, F.; Doranz, B.J.; Rucker, J.; Liesnard, C.; Farber, C.M.; Sarag

M.; Inai, T.; Rana, S.; Yi, Y.; Smyth, R.J.; Collman, R.G.; Doms, R.W.; Vassart, G.;

Nature 382, 722-725, 1996

A:Title: Resistance to HIV-1 infection in caucasian individuals bearing mutant allele

A:Reference number: S71808; MUID:96345670; PMID:8751444

A:Accession: S71808

A:Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 182-206;207-230 <SAM2>

A:Accession: A58834

A:Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-184, 'IKDSHLGAPAAACHGILLGNPKNSASVSK' <SAM3>

A:Cross-references: GB:X99393; NID:gl524062; PIDN:CAA67767.1; PID:gl524063

A:Note: This frameshift mutation results in a non-functional receptor but confers a de

nd may have had a selective advantage by conferring resistance to Yersinia plague infe

R:Combadiere, C.; Ahuja, S.K.; Tiffany, H.L.; Murphy, P.M.

J. Leukoc. Biol. 60, 147-152, 1996

A:Title: Cloning and functional expression of CC CKR5, a human monocyte CC chemokine 1

A:Reference number: A58832; MUID:96295970; PMID:8699119

A:Accession: A58832

A:Molecule type: mRNA

A:Residues: 1-352 <COM1>

A:Cross-references: GB:U57840; NID:gl502408; PIDN:AAB17071.1; PID:gl502409

A:Experimental source: clone 8, endotoxin-stimulated peripheral blood monocytes

R:Combadiere, C.

submitted to the EMBL Data Library, May 1996

A:Reference number: H01541

A:Accession: G02653

A:Status: translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-89, 'L', '91-352 <COM2>

A:Cross-references: EMBL:U57840

R:Raport, C.J.; Gosling, J.; Schweickart, V.L.; Gray, P.W.; Charo, I.F.

J. Biol. Chem. 271, 17161-17166, 1996

A:Title: Molecular cloning and functional characterization of a novel human CC chemoki

A:Reference number: A58833; MUID:96291862; PMID:8663314

A;Accession: A58833
A;Molecule type: mRNA
A;Residues: 1-352 <RAP>
C;Cross-references: GB:U54994; NID:G1457945; PIDN:AA05098.1; PID:G1457946
C;Comment: This is a receptor for chemokines MIP-1alpha (see PIR:A30574), MIP-1beta (see PIR:A30575) and dual-tropic strains of HIV-1 bind to a complex of chemokine receptors.
C;Genetics:
A;Gene: GDB:CMKBR5; CCR5; CCR-5; CC-CR-5; CCR5; ChemR13
A;Cross-references: GDB:1230510; OMIM:601373
A;Map position: 3p21-3p21
C;Function:
A;Description: G protein-coupled receptor for chemokines MIP-1alpha, MIP-1beta and RANTES.
A;Note: Probably acts to control granulocyte proliferation and differentiation.
C;Superfamily: vertebrate rhodopsin
C;Keywords: AIDS; G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane protein
F;32-56/Domain: transmembrane #status predicted <TM1>
F;67-87/Domain: transmembrane #status predicted <TM2>
F;103-124/Domain: transmembrane #status predicted <TM3>
F;142-166/Domain: transmembrane #status predicted <TM4>
F;193-218/Domain: transmembrane #status predicted <TM5>
F;236-257/Domain: transmembrane #status predicted <TM6>
F;285-300/Domain: transmembrane #status predicted <TM7>
F;20-269,101-178/Diulfide bonds: #status predicted
F;268/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;336,337,342/Binding site: phosphate (Ser) (covalent) #status predicted
F;340,343/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 47.6%; Score 379; DB 2; Length 352;
Best Local Similarity 35.0%; Pred. No. 2.5e-29;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

QY 1 MDYQVSPIDYDINVTSEPCQKINVKQIAA----- 30
DB 1 MDYQVSPIDYDINVTSEPCQKINVKQIAARLLPPLYSLVIFGFGVNMVLVILINCR 60
QY 31 -----YKGLC-----AAAQWDFGNTMCOHQRVHGHHHSYKCG----- 64
DB 61 LKSMTDIYLLNLATSDLFLLTPFWAHYAAQWDFGNTMCLLTGLYFIFGFGIF 118
QY 65 -----LCTRSQKGLHYTC 78
DB 119 ILLTIDRYLAHVHAFALKARTVFGVTSVITVVAVFASLPGLIFTRSQKGLHYTC 178
QY 79 SSSPPSYQYQFQKFNQTLKI-----HORVHG----- 105
DB 179 SSSPPSYQYQFQKFNQTLKIVGLVPLLVVICYSGLKTLRCRNEKRRHVALIF 238
QY 106 -----GGSYKGLC-----QFFGLNCCSSNRLDGHQV 135
DB 239 TIMIVYFLFWAPYINVLNTFQFFGLNCCSSNRLDQAMQV 281

RESULT 2
157505
zinc finger protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 05-Nov-1999
C;Accession: I57505
R;Ennault-Lange, M.; Kress, M.; Hamer, D.
Mol. Cell. Biol. 10, 418-421, 1990
A;Title: A gene that encodes a protein consisting solely of zinc finger domains is preferentially expressed in the developing mouse brain.
A;Reference number: I57505; MUID:90097859; PMID:2104662
A;Accession: I57505
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-201 <RES>
A;Cross-references: GB:M32057; NID:G199766; PIDN:AAA39729.1; PID:G199767
C;Genetics:
A;Gene: MOK2

Query Match 19.3%; Score 154; DB 2; Length 201;
Best Local Similarity 37.0%; Pred. No. 9.7e-08;
Matches 44; Conservative 13; Mismatches 44; Indels 18; Gaps 6;

QY 26 KQIAAYKCGLCAAAQWDFGNTMCOHQRVHGHHHSYKCGLCCTR--SQKGLHYTCSSHF- 82
DB 60 KPVAECGCGMSFSQR-----SNLHQRVH-TGERPYKCGCGKFSQSSNLIHRCCTHTG 114
QY 83 --PYSQYQFQKFN--QTLKIHQRVHGGSYKCGLCQEFFGLNCCSSNRLDGHQVH 136
DB 115 EKPQCYECGCGFSQSSDLRIHLRVHTGKPYHCGKCGQGF-----SOSKLLIHQRVH 168

RESULT 3
137570
zinc finger protein - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 01-Dec-2000
C;Accession: I37570
R;Abrink, M.; Aveskog, M.; Hellman, L.
DNA Cell Biol. 14, 125-136, 1995
A;Title: Isolation of cDNA clones for 42 different Kruppel-related zinc finger proteins
A;Reference number: I37566; MUID:95169271; PMID:7865130
A;Accession: I37570
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-693 <RES>
A;Cross-references: EMBL:X78927; NID:G498726; PIDN:CAA55527.1; PID:G498727
C;Genetics:
A;Gene: HZFA
C;Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology

Query Match 18.1%; Score 144; DB 2; Length 693;
Best Local Similarity 33.3%; Pred. No. 3e-06;
Matches 39; Conservative 15; Mismatches 33; Indels 30; Gaps 7;

QY 21 QKINVKQIAAYKCGLCAAA-QWDFGNTMCOHQRVHGHHHSYKCGLCCTRQKGLHYTC 79
DB 554 QKVHTG-KPYKCGCGKGFKWSLNDM--HQRVH-TGKPYTCGAC-----GRHFSQA 603
QY 80 SHFPYSQYQFQKFNQTLKIHQRVHGGSYKCGLCQEFFGLNCCSSNRLDGHQVH 136
DB 604 S-----SLQLHQSHTGKPYKCDVCGKVF-----SRSSQLQYHRRVH 641

RESULT 4
T46469
hypothetical protein DKFZp434G1930.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C;Accession: T46469
R;Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23034
A;Accession: T46469
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-319 <AAA>
A;Cross-references: EMBL:AL137483
A;Experimental source: adult testis; clone DKFZp434G1930
C;Genetics:
A;Note: DKFZp434G1930.1

Query Match 18.0%; Score 143.5; DB 2; Length 319;
Best Local Similarity 32.8%; Pred. No. 1.6e-06;
Matches 38; Conservative 15; Mismatches 44; Indels 19; Gaps 7;

QY 31 YKCGLCAAAQWDFGNTMCOHQRVHGHHHSYKCGLCCTRQKGLHYTCSSHF-----P 83
DB 9 YKCNCEKT-FSHRSLLSHQRIH-TGKPYKCNCEKAFSNS--STLNLHVRVHTGXP 64
QY 84 YSQYQFQKFN--QTLKIHQRVHGGSYKCGLCQEFFGLNCCSSNRLDGHQVH 136
DB 65 YRCGCGKAFSQCSLTIVHQRITHTGKLYKCGCEKAF---NCRA--KLHRHQRH 115

RESULT 9

148668
zinc finger protein 51 - mouse
N:Alternate names: finger protein zfc12
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 08-Dec-2000
C:Accession: I48668; A40984
R:Burke, P.S.; Don, J.; Wolgemuth, D.J.
Mamm. Genome 5, 387-389, 1994
A:Title: Zfp-51, a murine zinc finger encoding gene mapping to the t-complex region of chromosome 17
A:Reference number: I48668; MUID:94319090; PMID:8043957
A:Accession: I48668
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-710 <RES>
A:Cross-references: EMBL:X74855; NID:9488832; PIDN:CAA52847.1; PID:9488833
R:Crossley, P.H.; Little, P.F.R.
Proc. Natl. Acad. Sci. U.S.A. 88, 7923-7927, 1991
A:Title: A cluster of related zinc finger protein genes is deleted in the mouse embryonic stem cell line ES-1
A:Reference number: A40984; MUID:91376058; PMID:1680234
A:Accession: A40984
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 672-710 <CRO>
A:Cross-references: GB:M74235
C:Genetics:
A:Gene: Zfp-51
C:Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology
C:Keywords: DNA binding; zinc finger

Query Match 16.8%; Score 134; DB 2; Length 710;
Best Local Similarity 28.6%; Pred. No. 2.9e-05;
Matches 38; Conservative 20; Mismatches 39; Indels 36; Gaps 7;
C:Keywords: DNA binding; zinc finger

QY 9 IYDINVTYS-----EPCQKINVKQIAAYKCGLCGCAAAQDFGNTMCQHQHVHGHHSYK 63
DB 120 VYD-DYFSSAYSLMQPTAVIKEKPHQCKGK-----FNTSSSLTVHRIH-TKEKPYK 173
QY 64 GLCTRSQKGLHYTCSSHPYQYQFWKNFOTLKHQHVHGHHSYKCGLCQBFGLNLC 123
DB 174 SVCDKS-----FTQCTH-----LKHQHVHGHHSYKCGLCQBFGLNLC 208
QY 124 SSSNRLDGHQHVH 136
DB 209 VQLSALKSHQKHL 221

RESULT 10

S03679
finger protein (clone mkr5) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 01-Dec-2000
C:Accession: S03679
R:Chowdhury, K.; Rondewold, H.; Gruss, P.
Nucleic Acids Res. 16, 9995-10011, 1988
A:Title: Specific and ubiquitous expression of different Zn finger protein genes in the mouse embryo
A:Reference number: S03679; MUID:89057528; PMID:3143103
A:Accession: S03679
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-475 <CHO>
A:Note: the sequence in fig.2 is inconsistent with that shown in fig.1 in having 64-Asn, C;Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology

Query Match 16.8%; Score 133.5; DB 2; Length 475;
Best Local Similarity 27.8%; Pred. No. 2.2e-05;
Matches 40; Conservative 24; Mismatches 49; Indels 31; Gaps 8;
QY 15 YTSEPCOK-----INVQK-----IAAYKCGLCGCAAAQDFGNTMCQHQHVHGHHSYK 62
DB 223 YTCEVCHKSPRYGSSLTQVHRIHTGKPYECBICRKA-FSHASLTQHQVHS-GEKPFK 280

QY 63 CGLCTRSQKGLHYTCSSHF-----PYSQYQFWKNF-----QTLKHQHVHGHHSYKCG 112
DB 281 CKEGKAFQRIH--LASHWRIHTGKPECGCGSFSISSQLATHQRIHTGKPECK 338
QY 113 LCOEFFGLNCCSSNRLDGHQHVH 136
DB 339 VCRKAFRN-----IHLASHWRIH 357

RESULT 11

I48689
Gene NK10 protein - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 01-Dec-2000
C:Accession: I48689; S49078
R:Langre, R.; Christoph, A.; Thiesen, H.J.; Vopper, G.; Johnson, K.R.; Lemaire, L.; Ploma
DNA Cell Biol. 14, 971-981, 1995
A:Title: Developmentally regulated mouse gene NK10 encodes a xZinc Finger Repressor Prot
A:Reference number: I48689; MUID:96069544; PMID:7576184
A:Accession: I48689
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-636 <RES>
A:Cross-references: EMBL:X79828; NID:9506501; PIDN:CAA56225.1; PID:9506502
C:Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology

Query Match 16.8%; Score 133.5; DB 2; Length 636;
Best Local Similarity 31.6%; Pred. No. 2.9e-05;
Matches 37; Conservative 15; Mismatches 44; Indels 21; Gaps 6;

QY 31 YKGLCAAAQDFGNTMCQHQHVHGHHSYKCGLCGTR-----QKGLHYTCSSHF 82
DB 306 YQCSLCGKA-FQRSSSLVQHQRIH-TGEKPYQVLCGRSPRHSTSLTQHEVTH---SGEK 360
QY 83 PYSQYQFWKNF---OTLKHQHVHGHHSYKCGLCQBFGLNCCSSNRLDGHQHVH 136
DB 361 PPOCKEKGKAFRCSSSLVQHERHTGKPEKFCICGRAFG-----QSPS-YKGMRIH 412

RESULT 12

T14757
hypothetical protein DKFp572C163.1 - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 01-Dec-2000
C:Accession: T14757
R:Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, August 1999
A:Reference number: Z18181
A:Accession: T14757
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-701 <WAM>
A:Cross-references: EMBL:AL110217
A:Experimental source: adult subthalamic nucleus; clone DKFp572C163
C:Genetics:
A:Note: DKFp572C163.1
C:Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology

Query Match 16.8%; Score 133.5; DB 2; Length 701;
Best Local Similarity 23.4%; Pred. No. 3.2e-05;
Matches 41; Conservative 16; Mismatches 37; Indels 81; Gaps 8;
QY 31 YKGLCAAAQDFGNT---MCQHQHVH-----TCSSHPYQYQFWKNF-----GHHH-HS- 60
DB 468 YECNQK-----GKTFQRTLCQHQRIHTGKPYECNECGKTFADNSALRAHRIHTG 520
QY 61 ---YKGLCTRSQKGLHY-----TCSSHPYQYQFWKNF----- 93
DB 521 EKPTECNCGKTFKTSVHLRAHLTRSGEKPYECSECGKTFSEKSYVSAHQRVHTGKPY 580
QY 94 -----QTLKHQHVHGHHSYKCGLCQBFGLNCCSSNRLDGHQHVH 136
DB 581 ECVCKGKPAHNSLTQVHRIHTGKSEYECNDCGKTF-----SQKSHLSAQRIH 630

RESULT 13

A34612
zinc finger protein ZNF7 - human
N:Alternate names: zinc finger protein kox4
C:Species: Homo sapiens (man)
C:Date: 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 01-Dec-2000
C:Accession: A34612; A56409; S10421; I37972
R:Lania, L.; Monti, E.; Pannuti, A.; Pascucci, A.; Pengue, G.; Felicitello, I.; La Mantia
Genomics 6, 333-340, 1990
A:Title: cDNA isolation, expression analysis, and chromosomal localization of two human
A:Reference number: A34612; MUID:9016993; PMID:2106481
A:Accession: A34612
A:Molecule type: mRNA
A:Residues: 1-686 <LAN>
A:Cross-references: GB:M29580; NID:G340445; PIDN:AAA61313.1; PID:G340446; GB:J04751
R:Bray, P.; Lichter, P.; Thiesen, H.J.; Ward, D.C.; Dawid, I.B.
Proc. Natl. Acad. Sci. U.S.A. 88, 9563-9567, 1991
A:Title: Characterization and mapping of human genes encoding zinc finger proteins.
A:Reference number: A56409; MUID:9205132; PMID:1946370
A:Accession: A56409
A:Molecule type: DNA
A:Residues: 425-589 <BRA>
A:Cross-references: GB:M77170
R:Thiesen, H.J.
submitted to the EMBL Data Library, March 1990
A:Reference number: S10397
A:Accession: S10421
A:Molecule type: mRNA
A:Residues: 413-468 <THI>
A:Cross-references: EMBL:X52335; NID:G34165; PIDN:CAA36561.1; PID:G930095
R:Thiesen, H.J.
New Biol. 2, 363-374, 1990

A:Title: Multiple genes encoding zinc finger domains are expressed in human T cells.

A:Reference number: I37949; MUID:91145339; PMID:2288909
A:Accession: I37972
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 413-468 <RES>
A:Cross-references: EMBL:X52335; NID:G34165; PIDN:CAA36561.1; PID:G930095
C:Genetics:
A:Gene: GDB:ZNF7
A:Cross-references: GDB:I20509; OMIM:194531
A:Map position: 8q24.3-8q24.3
C:Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology
C:Keywords: DNA binding; zinc finger

Query Match 16.5%; Score 131.5; DB 2; Length 686;
Best Local Similarity 26.9%; Pred. No. 4.9e-05;
Matches 45; Conservative 11; Mismatches 46; Indels 65; Gaps 8;

QY 31 YKGLCAAQWDFG--NTMCQHORVGHGHHHSYKCGLCITRSQXEGHL----- 75

Db 441 YKCNCKTKA---FCGSSRLTJHQHRTHTGKPKFCDECGKFGVQSGSLTJHQHRTHTGKPK 496

QY 76 YTCSSSHFFPYQVFWNFKQ---TLKIHQHVHGGGSGYKCG 112

Db 497 YVNCDCGKAFSSQSSLIYHRIHKGKPYECLOCGKAFSMTQLTJHQHVHTGKPKYKCN 556

QY 113 LCOEFPGLNN-----CSS-----SNRLDGHQHVH 136

Db 557 ECGKAFSONTLFOHIIHAGVVKPYECGKAFSSSYLIEHRIH 603

RESULT 14

S70006
finger protein zfocl - guinea pig
C:Species: Cavia porcellus (guinea pig)
C:Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 05-Nov-1999
C:Accession: S70006
R:Rivolta, M.N.; Negrini, C.; Wilcox, E.R.
Biochim. Biophys. Acta 1306, 127-132, 1996

A:Title: A novel zinc finger gene preferentially expressed in the retina and the organ
A:Reference number: S70006; MUID:96221281; PMID:8634327

A:Accession: S70006
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-261 <RIV>

A:Cross-references: EMBL:L26335; NID:G1237277; PIDN:AAC42091.1; PID:G1237278
C:Keywords: zinc

Query Match 16.4%; Score 131; DB 2; Length 261;
Best Local Similarity 28.8%; Pred. No. 2.1e-05;
Matches 42; Conservative 16; Mismatches 44; Indels 44; Gaps 7;

QY 26 KQIAANYKCG-----LCAAQWDFGNT--MCOHORVGHGHHHSY 61

Db 85 KPYKCYECGKAFNPSPLQTHMRVHTGKPYCYECGCGRFSNSNLCMKHORVH-TGKPKF 143

QY 62 KCGLCITRS-----QKGLHVTCSHFFPYQVFWNFKQ---QTLKIHQHVHGGGSGYK 110

Db 144 KCBEGKAPHTSSLCMKHORVH---TGEKPYKCYECGKAFSSQSSSLCIHQHVHTGKPKYR 200

QY 111 CGLCQEFFGLNCSNRLDGHQHVH 136

Db 201 CCGCGKAF-----SQSSSLCIHQHVH 221

RESULT 15

JH0500

zinc finger protein 29 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Dec-2000

C:Accession: JH0500

R:Denny, P.; Ashworth, A.

Gene 106, 221-227, 1991

A:Title: A zinc finger protein-encoding gene expressed in the post-meiotic phase of spermatogenesis.

A:Reference number: JH0500; MUID:92039080; PMID:1937051

A:Accession: JH0500

A:Molecule type: mRNA

A:Residues: 1-614 <DEN>

A:Cross-references: GB:X55126; NID:G55470; PIDN:CAA38920.1; PID:G55471.

A:Experimental source: testis

C:Keywords: DNA binding; zinc finger

F:217-244/Region: zinc finger

F:245-272/Region: zinc finger

F:273-300/Region: zinc finger

F:301-328/Region: zinc finger

F:329-356/Region: zinc finger

F:357-384/Region: zinc finger

F:385-412/Region: zinc finger

F:413-440/Region: zinc finger

F:441-468/Region: zinc finger

F:469-496/Region: zinc finger

F:497-524/Region: zinc finger

F:525-552/Region: zinc finger

F:553-580/Region: zinc finger

F:581-608/Region: zinc finger

Query Match 16.2%; Score 129.5; DB 2; Length 614;
Best Local Similarity 30.3%; Pred. No. 6.8e-05;
Matches 37; Conservative 12; Mismatches 42; Indels 31; Gaps 6;

QY 31 YKGLCAAQWDFGNTMCQHORVGHGHHHSYKCGLCITRS-----QXEGHL----- 75

Db 418 YQCGEC-GKNFSRSSNLATHRTH-LVEKPYKGLCGKGSFSQSSSLIAHQHTGKPKYE 475

QY 76 -YTCSSSHFFPYQVFWNFKQFOTLKIHQHVHGGGSGYKCGLCQEFFGLNCSNRLDGHQHV 134

Db 476 CLTCGESFSWS-----SNLIKHORHTTGKPKYKCGCGKGF-----SQRSQLVVHOR 522

QY 135 VH 136

Db 523 TH 524

Wed Oct 22 06:03:23 2003

us-10-057-890a-10.rpr

Page 6

Search completed: October 21, 2003, 17:54:34
Job time : 23.922 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 17:40:14 ; Search time 13.0983 Seconds
(without alignments)
495.460 Million cell updates/sec

Title: US-10-057-890A-10

Perfect score: 797

Sequence: 1 MDYQVSSPIYDNYITSEPC.....GLNCCSSRLDGHORVHAA 138

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	379	47.6	352	1	CKR5_HUMAN
2	374	46.9	352	1	CKR5_PANTR
3	369	46.3	352	1	CKR5_PONPY
4	366	45.9	352	1	CKR5_PAPHA
5	365	45.8	352	1	CKR5_PYGBI
6	365	45.8	352	1	CKR5_PYGNE
7	365	45.8	352	1	CKR5_TRAFR
8	365	45.8	352	1	CKR5_TRAPH
9	364	45.7	352	1	CKR5_GORGO
10	363	45.5	352	1	CKR5_MACMU
11	361	45.3	352	1	CKR5_CERTO
12	359	45.0	352	1	CKR5_HYILE
13	357	44.8	352	1	CKR5_CERAE
14	272	34.1	354	1	CKR5_MOUSE
15	249	31.2	354	1	CKR5_RAT
16	156.5	19.6	474	1	2256_HUMAN
17	154	19.3	201	1	2239_MOUSE
18	151.5	19.0	645	1	2235_MOUSE
19	151.5	19.0	754	1	2287_HUMAN
20	150.5	18.9	501	1	2F96_MOUSE
21	146.5	18.4	458	1	2239_HUMAN
22	144.5	18.1	759	1	2287_MOUSE
23	144	18.1	698	1	2234_HUMAN
24	140.5	17.6	604	1	2305_HUMAN
25	140.5	17.6	744	1	2362_HUMAN
26	140.5	17.6	803	1	2226_HUMAN
27	140.5	17.6	913	1	2228_HUMAN
28	139.5	17.5	578	1	2192_HUMAN
29	139.5	17.5	614	1	2F28_MOUSE
30	139.5	17.5	682	1	2N45_HUMAN
31	139.5	17.5	751	1	2184_HUMAN
32	139	17.4	810	1	2334_HUMAN
33	138.5	17.4	223	1	GLI4_HUMAN

RESULT 1

ID	CKR5_HUMAN	STANDARD	PRT	352	AA
AC	P51681	O14692; O14693; O14696; O14697; O14698; O14699;			
AC	O14700; O14701; O14702; O14703; O14704; O14705; O14706; O14707;				
AC	O14708; O15538; Q9UPA4;				
DT	01-OCT-1996 (Rel. 34, Created)				
DT	01-OCT-1996 (Rel. 34, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CCR5)				
DE	(HIV-1 fusion co-receptor) (CHEMR13) (CD195 antigen).				
GN	CCR5 OR CMKR5.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=96241590; PubMed=8639485;				
RA	Samson M., Labbe O., Mollereau C., Vassart G., Parmentier M.;				
R*	"Molecular cloning and functional expression of a new human				
R*	CC-chemokine receptor gene.";				
RL	Biochemistry 35:3362-3367(1996).				
[2]					
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=96291862; PubMed=8663314;				
RA	Report C.J., Gosling J., Schweichart V.L., Gray P.W., Charo I.F.;				
RT	"Molecular cloning and functional characterization of a novel human				
RT	CC chemokine receptor (CCR5) for RANTES, MIP-1beta, and MIP-1alpha.";				
RL	J. Biol. Chem. 271:17161-17166(1996).				
[3]					
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=96295970; PubMed=8699119;				
RA	Combadiere C., Ahuja S.K., Tiffany H.L., Murphy P.M.;				
RT	"Cloning and functional expression of CC CKR5, a human monocyte CC				
RT	chemokine receptor selective for MIP-1(alpha), MIP-1(beta), and				
RT	RANTES.";				
RL	J. Leukoc. Biol. 60:147-152(1996).				
[4]					
RP	SEQUENCE FROM N.A.				
RX	McCombie W.R., Wilson R., Chen E., Gibbs R., Zuo L., Johnson D.,				
RA	Nhan M., Parnell L., Dedhia N., Ansari A., Mardis E., Schutz K.,				
RA	Goj L., la Baetide M., Kaplan N., Greco T., Touchman J.,				
RA	Muzny D., Chen C.N., Evans C., Fitzgerald M., See L.H., Tang M.,				
RA	Porcel B.M., Dragan Y., Giacalone J., Pae A., Powell E.,				
RA	Solinas K.A., Desliva U., Diaz-Perez S., Zhou X., Yu Y.,				
RA	Watanabe M., Doggett N., Garcia D., Sagripanti J.L.;				
RL	Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.				
[5]					
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=98001387; PubMed=9343222;				
RA	Kuhmann S.E., Platt E.J., Korak S.L., Kabat D.;				
RT	"Polymorphisms in the CCR5 genes of African green monkeys and mice				
RT	implicate specific amino acids in infections by simian and human				
RT	immunodeficiency viruses.";				
RL	J. Virol. 71:8642-8656(1997).				

ALIGNMENTS

34	138.5	17.4	670	1	ZN16_HUMAN	P17020 homo sapien
35	138.5	17.4	1029	1	Z197_HUMAN	O14709 homo sapien
36	136.5	17.1	348	1	Z134_HUMAN	P52741 homo sapien
37	136.5	17.1	659	1	Z304_HUMAN	Q9hcx3 homo sapien
38	135.5	17.0	643	1	ZN74_HUMAN	O16587 homo sapien
39	133.5	16.8	636	1	ZF90_MOUSE	O61967 mus musculu
40	132.5	16.6	642	1	ZN14_HUMAN	P17017 homo sapien
41	131.5	16.5	367	1	Z211_HUMAN	Q13398 homo sapien
42	131.5	16.5	446	1	ZN38_HUMAN	P17036 homo sapien
43	131.5	16.5	506	1	Z157_HUMAN	P51786 homo sapien
44	131.5	16.5	686	1	ZN07_HUMAN	P17097 homo sapien
45	131	16.4	739	1	YD49_HUMAN	Q9p2j8 homo sapien

Query Match 47.6%; Score 379; DB 1; Length 352;

Best Local Similarity 35.0%; Pred. No. 1.2e-29;

Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

Qy 1 MDYQVSSPIYDNYITSEPCQKINVKQIAA----- 30

Db 1 MDYQVSSPIYDNYITSEPCQKINVKQIAARLLPLYSLVFIFGVGMVLILINCKR 60

Qy 31 -----YKGLC-----AAQWDFGNTMCHQORVHGHHSYKCG--- 64

Db 61 LKSMTDIYLLNLALISDLFFLLTPFWAHYAQAQWDFGNTMC--QLLTGLYFIFGFSGIF 118

Qy 65 -----LCTRSQKGLHYTC 78

Db 119 IILLTIDRYLAVHVAFAKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKGLHYTC 178

Qy 79 SSHFPYSQYQFWKQNFOTLKI-----HORVHGG----- 105

Db 179 SSHFPYSQYQFWKQNFOTLKIIVLGLVPLLVVVCYSGILKTLRCRNEKKGHRVRLIF 238

Qy 106 -----GGSYKGLC-----QEPFGLNCCSSNRLDGHQV 135

Db 239 TIMIVYLFWAPYNIIVLLNTTQEFFGLNCCSSNRLDQAMQV 281

RESULT 2

CKR5_PANTR STANDARD; PRT; 352 AA.

AC P56440; 002778;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CCR5).

GN CCR5 OR CMKRS5.

OS Pan troglodytes (Chimpanzee).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

OX NCBI_TaxID=9598;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=97268687; PubMed=9108095;

RA Edinger A.L., Amedee A., Miller K., Doranz B.J., Endres M.,

RA Sharon M., Sanson M., Lu Z.H., Clements J.E., Murphy-Corb M.,

RA Peiper S.C., Parmentier M., Broder C.C., Doms R.W.;

RT "Differential utilization of CCR5 by macrophage and T cell tropic

RT simian immunodeficiency virus strains."

RL Proc. Natl. Acad. Sci. U.S.A. 94:4005-4010(1997).

RN [2]

RP SEQUENCE FROM N.A.

RX Zimmerman P.A., Buckler-White A., Alkhatib G.;

RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=98022612; PubMed=9359654;

RA Zhang L., Carruthers C.D., He T., Huang Y., Cao Y., Wang G., Hahn B.,

RA Ho D.D.;

RT "HIV type 1 subtypes, coreceptor usage, and CCR5 polymorphism."

RL AIDS Res. Hum. Retroviruses 13:1357-1366(1997).

RN [4]

RP SEQUENCE FROM N.A.

RX MEDLINE=97426118; PubMed=9282822;

RA Zacharova V., Zachar V., Goutin A.S.;

RT "Sequence of chemokine receptor gene CCR5 in chimpanzees, a natural

RT HIV type 1 host."

RL AIDS Res. Hum. Retroviruses 13:1159-1161(1997).

RN [5]

RP SEQUENCE FROM N.A.

RX MEDLINE=98090115; PubMed=9430250;

RA Pretet J.-L., Zerbib A., Girard M., Guillet J.-G., Butor C.;

RT "Chimpanzee CXCR4 and CCR5 act as coreceptors for HIV type 1."

RL AIDS Res. Hum. Retroviruses 13:1583-1587(1997).

CC -!- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,

CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY

CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE

CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR

CC DIFFERENTIATION.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AF003663; AAB62557.1; -.

CC EMBL; U94329; AAB58446.1; -.

CC EMBL; AF011542; AAB65742.1; -.

CC EMBL; U97666; AAC51670.1; -.

CC EMBL; AF011540; AAB65740.1; -.

CC EMBL; U89797; AAC03717.1; -.

CC Dr InterPro; IPR000276; GPCR_Rhodpsn.

CC Pfam; PF00001; 7tm 1; 1.

CC DR PRINTS; PS00237; GPCRHOPOPSN.

CC DR PROSITE; PS00237; G-PROTEIN RECF_F1_1; 1.

CC DR PROSITE; PS0262; G-PROTEIN RECF_F1_2; 1.

CC G-PROTEIN COUPLED RECEPTOR; Transmembrane; Glycoprotein; Sulfation.

CC KW DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).

CC FT DOMAIN 1 30

CC FT TRANSMEM 31 58 1 (POTENTIAL).

CC FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).

CC FT TRANSMEM 69 89 2 (POTENTIAL).

CC FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).

CC FT TRANSMEM 103 124 3 (POTENTIAL).

CC FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).

CC FT TRANSMEM 142 166 4 (POTENTIAL).

CC FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).

CC FT TRANSMEM 199 218 5 (POTENTIAL).

CC FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).

CC FT TRANSMEM 236 260 6 (POTENTIAL).

CC FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).

CC FT TRANSMEM 278 301 7 (POTENTIAL).

CC FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).

CC FT DISULFID 101 178 BY SIMILARITY.

CC FT MOD RES 3 3 SULFATION (BY SIMILARITY).

CC FT MOD RES 10 10 SULFATION (BY SIMILARITY).

CC FT MOD RES 14 14 SULFATION (BY SIMILARITY).

CC FT MOD RES 15 15 SULFATION (BY SIMILARITY).

CC FT CARBOHYD 268 268 N-LINKED (GLCNAC...) (POTENTIAL).

CC FT CONFLICT 123 123 T -> S (IN REF. 1).

CC SQ SEQUENCE 352 AA; 40539 MW; 4A3B698B80FE34C CRC64;

Query Match 46.9%; Score 374; DB 1; Length 352;

Best Local Similarity 34.6%; Pred. No. 3.8e-29;

Matches 98; Conservative 8; Mismatches 27; Indels 150; Gaps 7;

Qy 1 MDYQVSSPIYDNYITSEPCQKINVKQIAA----- 30

Db 1 MDYQVSSPIYDNYITSEPCQKINVKQIAARLLPLYSLVFIFGVGMVLILINCKR 60

Qy 31 -----YKGLC-----AAQWDFGNTMCHQORVHGHHSYKCG--- 64

Db 61 LKSMTDIYLLNLALISDLFFLLTPFWAHYAQAQWDFGNTMC--QLLTGLYFIFGFSGIF 118

Qy 65 -----LCTRSQKGLHYTC 78

Db 119 IILLTIDRYLAVHVAFAKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKGLHYTC 178

Qy 79 SSHFPYSQYQFWKQNFOTLKI-----HORVHGG----- 105

Db 179 SSHFPYSQYQFWKQNFOTLKIIVLGLVPLLVVVCYSGILKTLRCRNEKKGHRVRLIF 238

Qy 106 -----GGSYKGLC-----QEPFGLNCCSSNRLDGHQV 135


```
Db 239 TIMIVYFLWAPYNIIVLLNTFOEFFGLNCSNRLDQAMQV 281
RESULT 3
ID CCR5_PONPY STANDARD; PRT; 352 AA.
AC 097881;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE C-C chemokine receptor type 5 (C-CR-5) (CCR-5) (CCR5).
GN CCR5 OR CCR5R5.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99416438; PubMed=10486970;
RA Zhang Y.-W., Ryder O.A., Zhang Y.-P.;
RT "Sequence evolution of the CCR5 chemokine receptor gene in primates.";
RL Mol. Biol. Evol. 16:1145-1154(1999).
CC -!- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
CC DIFFERENTIATION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF075446; AAD19858.1; -
CC InterPro; IPR000276; GPCR_Rhodopsn.
CC Pfam; PF00001; 7tm_1; 1.
CC PRINTS; PR00237; GPCRHOODPSN.
CC PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
CC PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
KW DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 31 58
FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 69 89
FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 103 124
FT DOMAIN 125 141
FT TRANSMEM 142 166
FT DOMAIN 167 198
FT TRANSMEM 199 218
FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 236 260
FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 278 301
FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
FT DISULFID 101 178 BY SIMILARITY.
FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
FT SEQUENCE 352 AA; 40527 MW; F4E2F47135AF658A CRC64;
Query Match 46.3%; Score 369; DB 1; Length 352;
Best Local Similarity 34.3%; Pred. No. 1.2e-28;
Matches 97; Conservative 28; Mismatches 7; Gaps 7;
1 MDYQVSSPTDYIDYTSBPCQKINVKQIAARLLPFLYSLVFIQFVGVMVLILINCKR 60
31 -----YKGLC-----AAQWDFGNTMCQORVHGHHHSYKCG--- 64
61 LKSMTDIYLLNLALISDLFLLTVPFWAHYAAQWDFGNTMC--QLLTGLYFIQFSGIFF 118
65 -----LCTRSQKGLHYTC 78
119 IILLTIDRYLAIYHVAFLKARTVTIFGVVTSVITWVAVFASLPGIIFTRSQKGLHYTC 178
79 SSHFPYSQYQFKNPOTLKI-----HORVHG----- 105
179 SSHFPYSQYQFKNPOTLKIIVGLVLPVVMVICYSGILKTLACRNEKKRAVRLIF 238
106 -----GSGYKGLC-----QEPFGLNCSNRLDGHQRV 135
239 TIMIVYFLWAPYNIIVLLNTFOEFFGLNCSNRLDQAMQV 281
RESULT 4
ID CCR5_PAPHA STANDARD; PRT; 352 AA.
AC P56441;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE C-C chemokine receptor type 5 (C-CR-5) (CCR-5) (CCR5).
GN CCR5 OR CCR5R5.
OS Papio hamadryas (Hamadryas baboon), and
OS Papio anubis (Olive baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Papio.
OX NCBI_TaxID=9557, 9555;
RN [1]
RP SEQUENCE FROM N.A.
RX SPECIES=P.hamadryas;
RX MEDLINE=97268687; PubMed=9108095;
RA Edinger A.L., Amedee A., Miller K., Doranz B.J., Endres M.,
RA Sharron M., Samson M., Lu Z.H., Clements J.E., Murphy-Corb M.,
RA Peiper S.C., Parmentier M., Broder C.C., Doms R.W.;
RT "Differential utilization of CCR5 by macrophage and T cell tropic
RT simian immunodeficiency virus strains";
RL Proc. Natl. Acad. Sci. U.S.A. 94:4005-4010(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX SPECIES=P.hamadryas;
RX MEDLINE=99210133; PubMed=10195758;
RA Saksena N.K., Wang B., Novembre F.J., Bolton W., Smit T.K., Lal R.B.;
RT "Species-specific changes in the CCR5 gene from African and Asian
RT nonhuman primates.";
RL AIDS Res. Hum. Retroviruses 15:479-483(1999).
RN [3]
RP SEQUENCE FROM N.A.
RX SPECIES=P.anubis;
RA Benton P.A., Timanus D.K., Shearer M.H., Lee D.R., Kennedy R.C.;
RA Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
CC DIFFERENTIATION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
```

```

DR EMBL; AF005658; AAB62552.1; -
DR EMBL; AF105287; AAD20556.1; -
DR EMBL; AF105288; AAD20557.1; -
DR EMBL; AF105289; AAD20558.1; -
DR EMBL; AF105290; AAD20559.1; -
DR EMBL; AF023452; AAC63830.1; -
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; G-PROTEIN RECP_F1_1; 1.
DR PROSITE; PS0262; G-PROTEIN RECP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 31 58 1 (POTENTIAL).
FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 69 89 2 (POTENTIAL).
FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 103 124 3 (POTENTIAL).
FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 142 166 4 (POTENTIAL).
FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 199 218 5 (POTENTIAL).
FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 236 260 6 (POTENTIAL).
FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 278 301 7 (POTENTIAL).
FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
FT DISULFID 101 178 BY SIMILARITY.
FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
FT CARBOHYD 268 268 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 352 AA; 40489 MW; 5E1504A9BA1F8B2 CRC64;

Query Match 45.9%; Score 366; DB 1; Length 352;
Best Local Similarity 33.9%; Pred. No. 2.3e-28;
Matches 96; Conservative 9; Mismatches 28; Indels 150; Gaps 7;

QY 1 MDYQVSSPYDINYYTSEPCQKINVKQIAA----- 30
DB 1 MDYQVSSPYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGVGNILVVLINCKR 60
QY 31 -----YKGLC-----AAQWDFGNTMCQHQRVHGHHSYKCG--- 64
DB 61 LKSMTDIYLNLALISDLFLLTVPFWAHYAQAQWDFGNTWC--QLLTGLYFGTFSGIFF 118
QY 65 -----LCTRSQKEGLHYTC 78
DB 119 IILLTDRLYLAVHAFVAFKARTVFGVTSVITWVAVFASLPGIIFTRSQREGLYTC 178
QY 79 SSHFPYSQYQFKNFQTLKI-----HORVHGG----- 105
DB 179 SSHFPYSQYQFKNFQTLKIIVGLVPLLMVVICYSGLIKTLRCRNEKGRHRAVLIF 238
QY 106 -----CGSYKGLC-----QEFFGLNCCSSNRLDGHQRV 135
DB 239 TIMIVYFLWAPYNIIVLLNTFQEFFGLNCCSSNRLDQANQV 281

RESULT 5
ID_CKRS_PYGBI STANDARD; PRT; 352 AA.
AC 097880;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CKR5).
GN CKR5 OR CKR5.
OS Pygathrix bieti (Black snub-nosed monkey) (Rhinopithecus bieti).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Colobinae;
OC Pygathrix.

```

```

OX NCBI_TaxID=61621;
RN SEQUENCE FROM N.A.
RP MEDLINE=99416438; PubMed=10486970;
RX Zhang Y.-W., Ryder O.A., Zhang Y.-P.;
RT "Sequence evolution of the CCR5 chemokine receptor gene in primates.";
RL Mol. Biol. Evol. 16:1145-1154(1999).
CC -!- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
CC DIFFERENTIATION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF075445; AAD19857.1; -.
CC InterPro; IPR000276; GPCR_Rhodopsn.
CC Pfam; PF00001; 7tm.1; 1.
CC PRINTS; PR00237; GPCRHOOPS.
CC PROSITE; PS00237; G-PROTEIN RECP_F1_1; 1.
CC PROSITE; PS0262; G-PROTEIN RECP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 31 58 1 (POTENTIAL).
FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 69 89 2 (POTENTIAL).
FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 103 124 3 (POTENTIAL).
FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 142 166 4 (POTENTIAL).
FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 199 218 5 (POTENTIAL).
FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 236 260 6 (POTENTIAL).
FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 278 301 7 (POTENTIAL).
FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
FT DISULFID 101 178 BY SIMILARITY.
FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
SQ SEQUENCE 352 AA; 40585 MW; 4366F142730F938F CRC64;

Query Match 45.8%; Score 365; DB 1; Length 352;
Best Local Similarity 33.6%; Pred. No. 2.8e-28;
Matches 95; Conservative 10; Mismatches 28; Indels 150; Gaps 7;

QY 1 MDYQVSSPYDINYYTSEPCQKINVKQIAA----- 30
DB 1 MDYQVSSPYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGVGNILVVLINCKR 60
QY 31 -----YKGLC-----AAQWDFGNTMCQHQRVHGHHSYKCG--- 64
DB 61 LKSMTDIYLNLALISDLFLLTVPFWAHYAQAQWDFGNTWC--QLLTGLYFGTFSGIFF 118
QY 65 -----LCTRSQKEGLHYTC 78
DB 119 IILLTDRLYLAVHAFVAFKARTVFGVTSVITWVAVFASLPGIIFTRSQREGLYTC 178
QY 79 SSHFPYSQYQFKNFQTLKI-----HORVHGG----- 105
DB 179 SSHFPYSQYQFKNFQTLKIIVGLVPLLMVVICYSGLIKTLRCRNEKGRHRAVLIF 238
QY 106 -----CGSYKGLC-----QEFFGLNCCSSNRLDGHQRV 135

```

```

RESULT 7
CKR5_TRAFR
ID _CKR5_TRAFR STANDARD; PRT; 352 AA.
AC 097878;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CKR5).
GN CKR5 OR CWRK5.
OS Trachypithecus francoisi (Francois' langur) (Indochinese langur).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Colobinae;
OC Trachypithecus.
RX NCBI_TaxID=54180;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99416438; PubMed=10486970;
RA Zhang Y.-W., Ryder O.A., Zhang Y.-P.;
RT "Sequence evolution of the CKR5 chemokine receptor gene in primates." ;
RL Mol. Biol. Evol. 16:1145-1154(1999).
CC -!- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
CC DIFFERENTIATION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF075442; AAD19854.1; -
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC Pfam; PF00001; 7tm.1; 1.
CC PRINTS; PR00237; GPCRHHODPSN.
CC PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
CC PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 31 58 1 (POTENTIAL).
FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 69 89 2 (POTENTIAL).
FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 103 124 3 (POTENTIAL).
FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 142 166 4 (POTENTIAL).
FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).

```


RA Peiper S.C., Parmentier M., Broder C.C., Doms R.W.; Doms R.W.;
 RT "Differential utilization of CCR5 by macrophage and T cell tropic
 RL simian immunodeficiency virus strains.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:4005-4010(1997).
 CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 CC DIFFERENTIATION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF005659; AAB62553.1; -
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PRINTS; PR00237; GPCRHHODOPSN.
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
 DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
 FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 31 58 1 (POTENTIAL).
 FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 69 89 2 (POTENTIAL).
 FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 103 124 3 (POTENTIAL).
 FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 142 166 4 (POTENTIAL).
 FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 199 218 5 (POTENTIAL).
 FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 236 260 6 (POTENTIAL).
 FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 278 301 7 (POTENTIAL).
 FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
 FT DISULFD 101 178 BY SIMILARITY.
 FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
 FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
 FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
 FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
 SQ SEQUENCE 352 AA; 40515 MW; D0E6FCB9FE5EAC94 CRC64;
 Query Match 45.7%; Score 364; DB 1; Length 352;
 Best Local Similarity 33.9%; Pred. No. 3.5e-28;
 Matches 96; Conservative 8; Mismatches 29; Indels 150; Gaps 7;
 QY 1 MDYQVSSPIYDINYYTSPCKINVKQIAA----- 30
 DB 1 MDYQVSSPTDIDYTSBPCKTNVKGIAARLLPFLYSLVIFGFGVGNMLVILINCKR 60
 QY 31 -----YKGLG-----AAQNDFGNTMCOHORVHHHHSYKCG--- 64
 DB 61 LKSMTDIYLLNLAISDLFFLLTVFVFAHAAQNDFGNTMC--QLLTGLVYIGFSGIFF 118
 QY 65 -----LCTRSQKGLHYTC 78
 DB 119 ILLTIDRYLAIVHAFVALKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQKGLHYTC 178
 QY 79 SSHPPYSQYQWQKQPTLKI-----HQVHVG----- 105
 DB 179 SSHPPYSQYQWQKQPTLKIIVGLVPLLVVWVICYSGILKTLRCRNEKRGHRAVLIF 238
 QY 106 -----GGSYKGLG-----QEPFGLNCCSSNRLDGHORV 135
 DB 239 TIMIVYFLWAPYNIIVLLNTFQEPFGLNCCSSNRLDQAMQV 281

RESULT 10
 CKR5 MACMU
 ID CKR5 MACMU STANDARD; PRT; 352 AA.
 AC P79436; O02746;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (CCRS).
 GN CCR5 OR CMKR5.
 OS Macaca mulatta (Rhesus macaque),
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), and
 OS Macaca nemestrina (Pig-tailed macaque)
 OS Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 OC NCBI_TaxID=9544, 9541, 9545;
 RN [1] SEQUENCE FROM N.A.
 RP SPECIES=M.mulatta;
 RC MEDLINE=97184592; PubMed=9032394;
 RA Marcon L., Choe H., Martin K.A., Farzan M., Ponath P.D., Wu L.,
 RA Newman W., Gerard N., Gerard C., Sodroski J.;
 RT "Utilization of C-C chemokine receptor 5 by the envelope
 RT glycoproteins of a pathogenic simian immunodeficiency virus,
 RT SIVmac239.";
 RL J. Virol. 71:2522-2527 (1997).
 RN [2] SEQUENCE FROM N.A.
 RP SPECIES=M.mulatta; STRAIN=Indian macaque;
 RC MEDLINE=97213934; PubMed=9060623;
 RA Chen Z., Zhou P., Ho D.D., Landau N.R., Marx P.A.;
 RT "Genetically divergent strains of simian immunodeficiency virus use
 RT CCR5 as a coreceptor for entry.";
 RL J. Virol. 71:2705-2714 (1997).
 RN [3] SEQUENCE FROM N.A.
 RP SPECIES=M.mulatta;
 RC MEDLINE=21354176; PubMed=11461684;
 RA Margulies B.J., Hauer D.A., Clements J.E.;
 RT "Identification and comparison of eleven rhesus macaque chemokine
 RT receptors.";
 RL AIDS Res. Hum. Retroviruses 17:981-986 (2001).
 RN [4] SEQUENCE FROM N.A.
 RP SPECIES=M.mulatta, M.fascicularis, and M.nemestrina;
 RC MEDLINE=97268687; PubMed=9108095;
 RA Edinger A.L., Amedee A., Miller K., Doranz B.J., Endres M.,
 RA Sharon M., Samson M., Lu Z.H., Clements J.E., Murphy-Corb M.,
 RA Peiper S.C., Parmentier M., Broder C.C., Doms R.W.;
 RT "Differential utilization of CCR5 by macrophage and T cell tropic
 RT simian immunodeficiency virus strains.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:4005-4010 (1997).
 CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 CC DIFFERENTIATION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U77672; AAC51109.1; -
 DR EMBL; U77379; AAC51158.1; -
 DR EMBL; U96762; AAC34132.1; -
 DR EMBL; AF005660; AAB62554.1; -

```

DR EMBL; AF005661; AAB62555.1; --
DR EMBL; AF005662; AAB62556.1; --
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS00237; G-PROTEIN RECF F1_1; 1.
DR PROSITE; PS00237; G-PROTEIN RECF F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 31 58
FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 69 89
FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 103 124
FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 142 166
FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 199 218
FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 236 260
FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 278 301
FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
FT DISULFID 101 178 BY SIMILARITY.
FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 241 241 M -> I (IN REF. 3).
FT CONFLICT 292 292 I -> M (IN REF. 3).
SQ SEQUENCE 352 AA; 40507 MW; 58B96C8B5909FACB2 CRC64;

Query Match 45.5%; Score 363; DB 1; Length 352;
Best Local Similarity 33.6%; Pred. No. 4.4e-28;
Matches 95; Conservative 10; Mismatches 28; Indels 150; Gaps 7;

QY 1 MDYQVSSPYDINYYTSEPCKINVKQIAA----- 30
DB 1 MDYQVSSPYDINYYTSEPCKINVKQIAARLLPLYSLVFIFGFGVGNILVLLINCKR 60
QY 31 -----YKCGLC----- 64
DB 61 LKSMTDIYLLNLAISDLLELLTVPFVFAHYAAQWDFGNTMC--QLLTGIFYGFFSGIFF 118
QY 65 -----LCRSQKSGLHYTC 78
DB 119 IILLTDIYLAIVHAFVALKARTVTFGVTSVITWVAVFASLPGLIFTRSQREGLHYTC 178
QY 79 SSHFPYSQYQWKNFOTLKI-----HORVHGG----- 105
DB 179 SSHFPYSQYQWKNFOTLKNVILGLVPLLVVVICYSGLKTLKRCNEKRRHRAVRLIF 238
QY 106 -----GSKYKGLC-----QEFGLNCGSSNRLDGHQV 135
DB 239 TIMIVYFLWAPYINIVLLNTPOEFFGLNCGSSNRLDAQMV 281

RESULT 11
ID_CKRS_CERTO STANDARD; PRT; 352 AA.
AC O62743; O62744; O62745; O62746;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
GN C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CCR5).
OS Cercopithecus torquatus atys (Red-crowned mangabey) (Sooty mangabey).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9531;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=Isolate 079, 085, 087, and 089;
RX MEDLINE=98321155; PubMed=965699;
RA Chen Z., Gettie A., Ho D.D., Marx P.A.;
RT "Primary SIVem isolates use the CCR5 coreceptor from sooty mangabeys
naturally infected in west Africa: a comparison of coreceptor usage
of primary SIVem, HIV-2, and SIVmac.";
RL Virology 246:113-124(1998).
CC CC FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
DIFFERENTIATION.
CC CC SUBCELLULAR LOCATION: Integral membrane protein.
CC CC SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC CC EMBL; AF051902; AAC39830.1; --
CC CC EMBL; AF051903; AAC39831.1; --
CC CC EMBL; AF051904; AAC39832.1; --
CC CC EMBL; AF051905; AAC39833.1; --
CC CC InterPro; IPR000276; GPCR_Rhodopsn.
CC CC Pfam; PF00001; 7tm 1; 1.
CC CC PRINTS; PR00237; GPCR_Rhodopsn.
CC CC PROSITE; PS00237; G-PROTEIN RECF F1_1; 1.
CC CC PROSITE; PS0262; G-PROTEIN RECF F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 31 58
FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 69 89
FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 103 124
FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 142 166
FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 199 235
FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 236 260
FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 278 301
FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
FT DISULFID 101 178
FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
FT VARIANT 2 2 D -> E (IN ISOLATE 087).
FT VARIANT 3 3 Y -> D (IN ISOLATE 079).
FT VARIANT 25 25 V -> K (IN ISOLATE 087).
FT VARIANT 100 100 M -> K (IN ISOLATE 079).
FT VARIANT 107 107 L -> V (IN ISOLATE 089).
FT VARIANT 134 134 V -> G (IN ISOLATE 079).
FT VARIANT 146 146 V -> L (IN ISOLATES 085 AND 089).
FT VARIANT 340 340 T -> I (IN ISOLATE 079).
SQ SEQUENCE 352 AA; 40489 MW; 20A196E2D47E49CA CRC64;

Query Match 45.3%; Score 361; DB 1; Length 352;
Best Local Similarity 33.6%; Pred. No. 6.9e-28;
Matches 95; Conservative 9; Mismatches 29; Indels 150; Gaps 7;

QY 1 MDYQVSSPYDINYYTSEPCKINVKQIAA----- 30
DB 1 MDYQVSSPYDINYYTSEPCKINVKQIAARLLPLYSLVFIFGFGVGNILVLLINCKR 60
QY 31 -----YKCGLC----- 64
DB 61 LKSMTDIYLLNLAISDLLELLTVPFVFAHYAAQWDFGNTMC--QLLTGIFYGFFSGIFF 118
QY 65 -----LCRSQKSGLHYTC 78
DB 119 IILLTDIYLAIVHAFVALKARTVTFGVTSVITWVAVFASLPGLIFTRSQREGLHYTC 178
QY 79 SSHFPYSQYQWKNFOTLKI-----HORVHGG----- 105
DB 179 SSHFPYSQYQWKNFOTLKNVILGLVPLLVVVICYSGLKTLKRCNEKRRHRAVRLIF 238
QY 106 -----GSKYKGLC-----QEFGLNCGSSNRLDGHQV 135
DB 239 TIMIVYFLWAPYINIVLLNTPOEFFGLNCGSSNRLDAQMV 281

```


entities requires a license agreement (See <http://www.isb-sib.ch/announce/>) or send an email to license@isb-sib.ch.

```
CC -----
CC EMBL; U83324; AAC51795.1; --
CC EMBL; U83325; AAC51796.1; --
CC EMBL; AB015944; BAA31328.1; --
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC Pfam; PF00001; 7tm.1; 1.
CC PRINTS; PR00237; GPCRHOOPS.
CC PROSITE; PS00237; G_PROTEIN_RECEP_FL_1; 1.
CC PROSITE; PS00242; G_PROTEIN_RECEP_FL_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation;
KW Polymorphism.
CC DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 31 58 1 (POTENTIAL).
CC FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 69 89 2 (POTENTIAL).
CC FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 103 124 3 (POTENTIAL).
CC FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 142 166 4 (POTENTIAL).
CC FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 199 218 5 (POTENTIAL).
CC FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 236 260 6 (POTENTIAL).
CC FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 278 301 7 (POTENTIAL).
CC FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
CC FT DISULFID 101 178 BY SIMILARITY.
CC FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
CC FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
CC FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
CC FT VARIANT 14 14 N -> Y.
CC FT VARIANT 352 352 F -> L.
CC SQ SEQUENCE 352 AA; 40561 MW; 7F52E690C72EC29A CRC64;

Query Match 44.8%; Score 357; DB 1; Length 352;
Best Local Similarity 33.6%; Pred. No. 1.7e-27;
Matches 95; Conservative 9; Mismatches 29; Indels 150; Gaps 7;

QY 1 MDYQSSPIYDINVTSEPCOKINVKQIAA----- 30
Db 1 MDYQSSPIYDINVTSEPCOKINVKQIAARLLPPLYSLVIFGFGVGNILVLLINCKR 60
QY 31 -----YKCGLC-----AAQWDFGNTWCQQRVGHGHHHSYKCG--- 64
Db 61 LKSMTDIYLLNLALISDLLFLTVPFWAHYAAQWDFGNTWC---QLLTGLYPIFGFSGIFF 118
QY 65 -----LCTRSQKEGLHYTC 78
Db 119 IILLTDIYLAIVHAVFALKARTVTFGWVTSVITWVAVFASLPRIIFTRSORQEGLYHYTC 178
QY 79 SSHFPYSQYQFWKQNFOTLKI-----HORVHGG----- 105
Db 179 SSHFPYSQYQFWKQNFOTLKIIVGLVPLLVWVICYSGILKTLRCNKKHRAVRLIF 238
QY 106 -----GGSYKCGLC-----QEFGLNCSSNRLDGHQRV 135
Db 239 TIMIVVFLFWAPYINIVLLNTFQEFFGLNCSSNRLDAQOV 281

RESULT 14
CKR5_MOUSE
ID CKR5_MOUSE STANDARD; PRT; 354 AA.
AC P51682; O35313; O35891; P97308; P97405; Q61867;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (MIP-1
DE alpha receptor).
GN CKR5 OR CMKBR5.
OS Mus musculus (Mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ; TISSUE=Spleen;
RX MEDLINE=96205938; PubMed=8631787;
RA Boxing L., Gosling J., Montecarlo F.S., Lusis A.J., Tsou C.-L.,
RA Charo I.F.;
RT "Molecular cloning and functional expression of murine JE (monocyte
RT chemoattractant protein 1) and murine macrophage inflammatory protein
RT alpha receptors: evidence for two closely linked C-C chemokine
RT receptors on chromosome 9.";
RL J. Biol. Chem. 271:7551-7558(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CS7BL/6 X CBA; TISSUE=Thymus;
RX MEDLINE=96278910; PubMed=8662890;
RA Meyer A., Coyle A.J., Proudfoot A.E.I., Wells T.N.C., Power C.A.;
RT "Cloning and characterization of a novel murine macrophage
RT inflammatory protein-1 alpha receptor.";
RL J. Biol. Chem. 271:14445-14451(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Ola;
RX Kuziel W.A., Beck M.A., Dawson T.C., Maeda N.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=CS7BL/6, and NIH Swiss; TISSUE=Kidney, Liver, and Spleen;
RX MEDLINE=98001387; PubMed=934322;
RA Kuhlmann S.E., Platt E.J., Kozak S.L., Kabat D.;
RT "Polymorphisms in the CCR5 genes of African green monkeys and mice
RT implicate specific amino acids in infections by simian and human
RT immunodeficiency viruses.";
RL J. Virol. 71:8642-8656(1997).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RX MEDLINE=97404635; PubMed=9261347;
RA Doranz B.J., Lu Z.H., Rucker J., Zhang T.Y., Sharron M., Cen Y.H.,
RA Wang Z.X., Guo H.H., Du J.G., Accavitti M.A., Doms R.W., Peiper S.C.;
RT "Two distinct CCR5 domains can mediate coreceptor usage by human
RT immunodeficiency virus type 1.";
RL J. Virol. 71:6305-6314(1997).
RN [6]
RP SEQUENCE FROM N.A.
RA Guo B., Kuno K., Harada A., Matsushima K.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: DETECTED IN MONOCYTE/MACROPHAGE CELL LINES,
CC BUT NOT IN NONHEMATOPOIETIC CELL LINES.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U47036; AAC52454.1; --
CC EMBL; X94151; CAA63867.1; --
CC EMBL; U68565; AAB37273.1; --
CC EMBL; U83327; AAC53386.1; --
CC EMBL; AF022990; AAC53389.1; --
CC EMBL; AF019772; AAB71183.1; --
CC EMBL; D83648; BAA12024.1; --
CC MGD; MGI:107182; Ccr5.
CC GO; GO:0016493; F:C-C chemokine receptor activity; IDA.
```


DR GO; GO:0006952; P:defense response; IMP.
DR InterPro; IPR000276; GPCR Rhodpsn.

DR	GO: 0006952; P: defense response; IMP.	
DR	InterPro; IPR000276; GPCR_Rhodpsn.	
DR	Pfam; PF00001; 7tm 1; 1.	
DR	PROSITE; PRO0237; GPCRHOOPS.	
DR	PROSITE; PS00237; G PROTEIN RECP_F1_1; 1.	
DR	PROSITE; PS00237; G PROTEIN RECP_F2_1; 1.	
DR	PROSITE; PS0362; G-PROTEIN RECP_F1_1.	
KW	G-protein coupled receptor; Transmembrane; Glycoprotein; Polymorphism.	
FT	DOMAIN 1 32	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 33 60	1 (POTENTIAL).
FT	DOMAIN 61 70	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 71 91	2 (POTENTIAL).
FT	DOMAIN 92 104	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 105 126	3 (POTENTIAL).
FT	DOMAIN 127 143	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 144 168	4 (POTENTIAL).
FT	DOMAIN 169 200	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 201 220	5 (POTENTIAL).
FT	DOMAIN 221 237	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 238 262	6 (POTENTIAL).
FT	DOMAIN 263 279	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 280 303	7 (POTENTIAL).
FT	DOMAIN 304 354	CYTOPLASMIC (POTENTIAL).
FT	DISULFID 103 180	BY SIMILARITY.
FT	CARBOHYD 270 270	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	VARIANT 11 11	I -> S.
FT	VARIANT 62 62	K -> R.
FT	VARIANT 66 66	V -> M.
FT	VARIANT 97 97	I -> V.
FT	VARIANT 109 109	V -> L.
FT	VARIANT 156 156	V -> A.
FT	VARIANT 160 160	F -> S.
FT	VARIANT 185 185	P -> L.
FT	VARIANT 213 213	I -> V.
FT	VARIANT 318 318	I -> M.
FT	VARIANT 337 337	V -> A.
FT	CONFLICT 3 3	F -> L (IN REF. 2).
FT	CONFLICT 80 80	L -> P (IN REF. 2).
FT	CONFLICT 145 145	N -> I (IN REF. 5).
FT	CONFLICT 190 190	H -> Y (IN REF. 3).
FT	CONFLICT 208 208	P -> S (IN REF. 1).
SQ	SEQUENCE 354 AA; 40863 MW; BBA6B942E88F9CF0 CRC64;	

Query Match 34.1%; Score 272; DB 1; Length 354;

Query Match 34.1%; Score 272; DB 1; Length 354;
Best Local Similarity 27.9%; Pred. No. 2.8e-19;
Matches 80; Conservative 12; Mismatches 29; Indels 166; Gaps 8;

QY	1	MDYQVSSP--IYDINYYSBFCOKINVKQIAA-----	30
Db	1	MDFOGSVPTIYDIDYGMSPCKINVKQIAAQLLPYLSLVTFIFGVGNMVMVFLILISC	60
QY	31	-----YKCGLC-----AAQWDFGNTMQHORVGHGHHHSYKCG--	64
Db	61	KKLKSVTDIYLLNLALISDLLFLTLPPWHAAYANEMIFGNIMC--KVFTGVVHIGYFGGI	118
QY	65	-----LCTRSQKEGLHY	76
Db	119	FFIILLTIDRYLAIHVAFALKVTFNVFVITSVTVVWVAFSLPELIIFTRSQEGPHY	178
QY	77	TCSSHPYPYQOYFQWKNFOTLK-----IHORVHGCGGSKYC-----	111
Db	179	TCSPPHPTQYFHWKSFQTLKMWLSLILPLLWMIICYSGILHT-----LPRCNEKK	231
QY	112	-----GLCQEPFGLNCCSSNRLD	130
Db	232	RHRVRLIEAIVMIVYFLEWTPNIVILLTTTFOBFPGNLNCCSSNRLD	278

RESULT 15

CKR5_RAT	STANDARD;	PRT;	354 AA.
ID CKR5_RAT			

AC 008556;
DT 01-NOV-1997 (Revised) Created)

DT	01-NOV-1997	{REL. 35, Created}
DT	01-NOV-1997	{REL. 35, Last sequence update}

30-MAY-2000 (Rel. 39, Last annotation update)
C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (MIP-1
alpha receptor).
CCRS OR CWKBR5.
Rattus norvegicus (Rat).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
STRAIN=Mistar; TISSUE=Brain;
MEDLINE=98334064; PubMed=9670989;
Spieles O., Gourmalan N., Boddeke H.W.G.M., Sauter A., Fiebich B.L.,
Berger M., Gebicke-Haerter P.J.;
"Cloning of rat HIV-1-chemokine coreceptor CCR5 from microglia and
upregulation of its mRNA in ischemic and endotoxemic rat brain.";
J. Neurosci. Res. 53:16-28(1998).
[2]
SEQUENCE FROM N.A.
STRAIN=Sprague-Dawley;
MEDLINE=98318173; PubMed=9655467;
Jiang Y., Salafraanca M.N., Adhikari S., Xia Y., Feng L., Sonntag M.K.,
Defebvre C.M., Pennell N.A., Sireit W.J., Harrison J.K.;
"Chemokine receptor expression in cultured glia and rat experimental
allergic encephalomyelitis";
J. Neuroimmunol. 86:1-12(1998).
-!- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL.
-!- SUBCELLULAR LOCATION: Integral membrane protein.
-!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to licenseseis@sib.ch).

CC
DR EMBL; Y12009; CAA72737.1; -;

DR EMBL; U77350; AAC03243.1; -;
DR InterPro; IPR000276; GPCR Rhodpsn.

DR Pfam; PF00001; 7tm_1; 1.

DR PRINTS; PR00237; GPCRRHODPSN.

DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.

DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane;

FT	DOMAIN	1	32	EXTRACELLULAR (POTENTIAL)	TRANSMEMBRANE; GLYCOPROTEIN.
KW	G-protein coupled receptor				

FT	TRANSMEM	33	60	1 (POTENTIAL).

FT	DOMAIN	61	70	CYTOPLASMIC (POTENTIAL).
100	100	100	100	100
90	90	90	90	90
80	80	80	80	80
70	70	70	70	70
60	60	60	60	60
50	50	50	50	50
40	40	40	40	40
30	30	30	30	30
20	20	20	20	20
10	10	10	10	10
0	0	0	0	0

FT	71	91	2 (POTENTIAL).
TRANSMEM			
DOMAIN	92	104	EXTRACELLULAR (POTENTIAL)
ET			

FT	105	126	3 (POTENTIAL).
DOMAIN	92	104	EXTRACELLULAR (POTENTIAL).

FT	DOMAIN	127	143	CYTOPLASMIC (POTENTIAL).
FT				

FT	TRANSMEM	144	168	4 (POTENTIAL).
----	----------	-----	-----	----------------

FT	DOMAIN	169	200	EXTRACELLULAR (POTENTIAL).
cm				

FT	DOMAIN	201	220	5 (POTENTIAL)	CYTOSOLIC / POTENTIAL
FT	DOMAIN	201	220	5 (POTENTIAL)	CYTOSOLIC / POTENTIAL

DOMAIN	221	237	CYTOPLASMIC (POTENTIAL).
FT		262	6 (POTENTIAL)
TRANSMEM	238		

TRANSMEM	258	8 (POTENTIAL)/.
FT DOMAIN	263	EXTRACELLULAR (POTENTIAL).
FT	279	

FT	TRANSMEM	280	303	7 (POTENTIAL).

	FT	DOMAIN	304	354	CYTOPLASMIC (POTENTIAL).
FT					
DOMAIN					
304					
354					
CYTOPLASMIC (POTENTIAL)					

FT	DISULFID	103	180	BY SIMILARITY.
FT	CARBOHYD	270	270	N-1 LINKED / CYCLOCAC

FT	270	N-LINKED (GLCNAC. :.) (POTENTIAL).
CARBOHYD	270	
SEQUENCE	354 AA:	41030 MW: 77EDB368AA4C868D CRC64:

Query Match 31.2%: Score 249: DB 1: length 354:

Query Match 31.2%; Score 243; DB 1; Length 334;
Best Local Similarity 27.0%; Pred. No. 4,7e-17;

Matches	77;	Conservative	16;	Mismatches	40;	Indels	152;	Gaps	8;
---------	-----	--------------	-----	------------	-----	--------	------	------	----

QY 1 MDYQVSSP--IYDINYTSEPCQKINVKQIAA-----30

Db 1 MDFQGSIPTYYIDYMSAPCQKVNKQIAAQLLPPLYSLVIFGFVGNMVFLLISC 60
Qy 31 -----YKGLC-----RAAOWDFGNTMCOHQRVHGHHSYKCG- 64
Db 61 KKLKSMTDIYLFNLAISDLLLTLFPWAHYAANEWFVGNIMC--KLFTGIYHIGYFGI 118
Qy 65 -----LCTRSQKEGLHY 76
Db 119 FFIILLTIDRY:LAIVHAVFAIKARTVNEGVITSVVTVWVAVFVSLPEIIFMRSQKEGSHY 178
Qy 77 TCSSHPYQYOFKWFQTLK-----IHORVHGG--GGSYKC----- 111
Db 179 TCSPHFLHQYRFWKHFQTLKMWILSLILPLLVVICYSGLINTLFCRNEKXGHRVRL 238
Qy 112 -----GLCOEFFGLNCCSSNRLDGHRV 135
Db 239 IFAIMIVYFLFWTPYINIVLLTTFOEYFGLNCCSSNRLDQMQRV 283

Search completed: October 21, 2003, 17:51:29
Job time : 14.0983 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 17:47:34 ; Search time 57.0712 Seconds
(without alignments)
623.980 Million cell updates/sec

Title: US-10-057-890A-10

Perfect score: 797

Sequence: 1 MDYQVSSPIYDINYTSEPC.....GLNCCSSNRLDGHQHVAA 138

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phase.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	373	46.8	352	6 018772	018772 pan troglod
2	370	46.4	344	6 09TOR8	09tor8 cercocobus
3	370	46.4	344	6 077833	077833 cercocobus
4	369	46.3	352	6 095NC5	095nc5 hylobates s
5	369	46.3	352	6 09TV50	09tv50 pan troglod
6	367	46.0	352	6 095NC3	095nc3 miopithecus
7	367	46.0	352	6 018771	018771 pan troglod
8	366	45.9	352	6 09TSTK1	09tsk1 cercopithec
9	366	45.9	352	6 09TV49	09rv49 cercocobus
10	366	45.9	352	6 09XT13	09xt13 papio anubi
11	366	45.9	352	6 09TV45	09tv45 cercopithec
12	366	45.9	352	6 095NE8	095ne8 cercopithec
13	365	45.8	352	6 09XT14	09xt14 colobus gue
14	365	45.8	352	6 095NC6	095nc6 trachypithe
15	365	45.8	352	6 095NC8	095nc8 colobus pol
16	364	45.7	352	6 09XS99	09xs99 gorilla gor

```

17 363 45.5 352 6 097975 097975 macaca arct
18 363 45.5 352 6 09MZA2 09mza2 cercopithec
19 362 45.4 352 6 095NC1 095nc1 theropithec
20 362 45.4 352 6 095NC0 095nc0 hylobates m
21 362 45.4 352 6 09TV46 09tv46 cercopithec
22 362 45.4 352 6 097962 097962 pygathrix a
23 361 45.3 352 6 095NC7 095nc7 nasalis lar
24 361 45.3 352 6 09TV42 09tv42 cercopithec
25 361 45.3 352 6 077776 077776 cercocobus
26 361 45.3 352 6 095NE1 095ne1 cercocobus
27 361 45.3 352 6 09TOX0 09tox0 cercopithec
28 359 45.0 352 6 09TV43 09tv43 cercopithec
29 358 44.9 352 6 018770 018770 pan troglod
30 357 44.8 352 6 09TV47 09tv47 cercopithec
31 357 44.8 352 6 09BGN5 09bgn5 cercopithec
32 357 44.8 352 6 09XT12 09xt12 cercopithec
33 357 44.8 352 6 09TSQ7 09tsq7 cercopithec
34 356 44.7 352 6 095ND1 095nd1 mandrillus
35 356 44.7 352 6 095ND1 095nd1 mandrillus
36 355 44.5 352 6 09XT76 09xt76 cercopithec
37 355 44.5 352 6 09TV44 09tv44 cercopithec
38 355 44.5 352 6 09XS35 09xs35 macaca neme
39 355 44.5 352 6 09MZA3 09mza3 hylobates a
40 355 44.5 352 6 09TV93 09tv93 macaca arct
41 355 44.5 352 6 095ND0 095nd0 erythrocebu
42 346 43.4 352 6 09BGN6 09bgn6 cercopithec
43 345 43.3 339 4 09UN25 09un25 homo sapien
44 344 43.2 339 4 09UN23 09un23 homo sapien
45 344 43.2 339 4 09UN27 09un27 homo sapien

```

ALIGNMENTS

RESULT 1

```

018772 ID 018772 PRELIMINARY; PRT; 352 AA.
AC 018772;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CCR5 receptor (Fragment).
GN CCR5.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHCCRS-142a;
RA Zhang L., Caruthers C.D., He T., Huang Y., Cao Y., Wang G., Hahn B.,
RA Ho D.D.;
RT "HIV-1 subtypes, co-receptor usage, and CCR5 polymorphism.";
RL AIDS Res Hum Retroviruses 0:0-0(1997).
DR EMBL; AF011541; AAB65741.1; -.
DR InterPro; IPR000276; GPCR_Rhodopsin.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER 352 352
SQ SEQUENCE 352 AA; 40598 MW; 39688FA7004C952F CRC64;

```

Query Match 46.8%; Score 373; DB 6; Length 352;

Best Local Similarity 34.6%; Pred. No. 1.3e-33;

Matches 98; Conservative 8; Mismatches 27; Indels 150; Gaps 7;

QY 1 MDYQVSSPIYDINYTSEPCQKINVKQIAA----- 30

D5 1 MDYQVSSPIYDIDYTTSEPCQKINVKQIAARLLPPLYSLVIFGFVDNMLVILLINCKR 60

QY 31 -----YKGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64

```

Db      61 LKSMTDIYLLNLAIISDLFFLLTVPFWAHYAAQWDFGNTMC--QLLTGLYFIFFSGIFF 118
QY      65 -----YKGLC-----AAQWDFGNTMCQHQHVHGHHSYKCG--- 64
Db      119 ILLTIDRYLAIVHAVFALKARTVFGVTSVITVVAVFASLPGLIIFTRSQEGLHYTC 178
QY      79 SSHFPYQYQWKNFQTLKI-----LCTRSQKGLHYTC----- 105
Db      179 SSHFPYQYQWKNFQTLKIIVILGLVPLLVVVICYSGLKTLRLCRNEKGRHRAVRLIF 238
QY      106 -----GGSYKGLC-----QEFFGLNCCSSNRLDGHQV 135
Db      239 TIMIVYFLWAPYDIVLLNTFQEFFGLNCCSSNRLDQAMQV 281

RESULT 2
Q9TOR8 PRELIMINARY; PRT; 344 AA.
AC Q9TOR8;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DE 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Chemokine receptor CCR5.
GN CCR5.
OS Cercocobus torquatus torquatus, and
OS Cercocobus torquatus atys (Red-crowned mangabey) (Sooty mangabey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Cercocobus.
OX NCBI_TaxID=81944, 9531;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=C.torquatus torquatus; STRAIN=1049, and 997;
RA Chen Z., Kwon D., Jin Z., Monard S., Telfer P., Jones M., Lu C.,
RA Aguilar R., Ho D.D., Marx P.A.;
RT "Natural infection of a homozygous delta 24 CCR5 red-capped mangabey
RT with a R2b-tropic simian immunodeficiency virus.";
RL J. Exp. Med. 0:0-0(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=C.torquatus atys;
RA Palacios E., Digilio L., McClure H.M., Chen Z., Marx P.A.,
RA Goldsmith M.A., Grant R.M.;
RT "Parallel evolution of CCR5-null phenotypes in humans and in a natural
RT host of simian immunodeficiency virus.";
RL Curr. Biol. 0:0-0(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=C.torquatus torquatus; STRAIN=RCM411;
RA Beer B.E., Kuiken C.L., Toozé Z., Foley B.T., Goeken R.M., Brown C.R.,
RA St. Claire M., Hirsch V.M.;
RT "Characterization of novel simian immunodeficiency viruses from
RT red-capped mangabays from Nigeria (SIVrcmNgM411 and NgD409).";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF094753; AAC62474.1; -
DR EMBL; AF079473; AAC31194.1; -
DR EMBL; AF084003; AAC62471.1; -
DR EMBL; AF349683; AAC69685.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PRO0237; GPCRHHODPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 344 AA; 39592 MW; E15F5F601191A4D1 CRC64;

Query Match 46.4%; Score 370; DB 6; Length 344;
Best Local Similarity 34.9%; Pred. No. 2.8e-33;
Matches 96; Conservative 9; Mismatches 28; Indels 142; Gaps 7;

QY      1 MDYQVSSPIYDINYTSEPCQKINVKQIAA----- 30

```

```

Db      31 MDYQVSSPTDYDITSEPCQKINVKQIAARLLPPLYSLVFIQFVGNILVLLINCKR 60
QY      61 LKSMTDIYLLNLAIISDLFFLLTVPFWAHYAAQWDFGNTMC--QLLTGLYFIFFSGIFF 118
QY      65 -----YKGLC-----AAQWDFGNTMCQHQHVHGHHSYKCG--- 64
Db      119 ILLTIDRYLAIVHAVFALKARTVFGVTSVITVVAVFASLPGLIIFTRSQEGLHYTC 178
QY      79 SSHFPYQYQWKNFQTLKI-----LCTRSQKGLHYTC----- 105
Db      179 SSHFPYQYQWKNFQTLKIIVILGLVPLLVVVICYSGLKTLRLCRNEKGRHRAVRLIF 238
QY      106 -----GGSYKGLC-----QEFFGLNCCSSNRLDGHQV 135
Db      239 TIMIVYFLWAPYDIVLLNTFQEFFGLNCCSSNRLDQAMQV 281

RESULT 3
Q77833 PRELIMINARY; PRT; 344 AA.
AC Q77833;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DE 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Chemokine receptor CCR5.
GN CCR5.
OS Cercocobus torquatus torquatus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Cercocobus.
OX NCBI_TaxID=81944;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1208, and 009;
RA Chen Z., Kwon D., Jin Z., Monard S., Telfer P., Jones M., Lu C.,
RA Aguilar R., Ho D.D., Marx P.A.;
RT "Natural infection of a homozygous delta 24 CCR5 red-capped mangabey
RT with a R2b-tropic simian immunodeficiency virus.";
RL J. Exp. Med. 0:0-0(1998).
DR EMBL; AF094752; AAC62473.1; -
DR EMBL; AF084002; AAC62470.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PRO0237; GPCRHHODPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
SQ SEQUENCE 344 AA; 39578 MW; 0CEC05E47C2F6DE6 CRC64;

Query Match 46.4%; Score 370; DB 6; Length 344;
Best Local Similarity 34.9%; Pred. No. 2.8e-33;
Matches 96; Conservative 9; Mismatches 28; Indels 142; Gaps 7;

QY      1 MDYQVSSPIYDINYTSEPCQKINVKQIAA----- 30
Db      1 MDYQVSSPTDYDITSEPCQKINVKQIAARLLPPLYSLVFVFGVGNILVLLINCKR 60
QY      31 MDYQVSSPTDYDITSEPCQKINVKQIAA-----AAQWDFGNTMCQHQHVHGHHSYKCG--- 64
Db      61 LKSMTDIYLLNLAIISDLFFLLTVPFWAHYAAQWDFGNTMC--QLLTGLYFIFFSGIFF 118
QY      65 -----YKGLC-----AAQWDFGNTMCQHQHVHGHHSYKCG--- 64
Db      119 ILLTIDRYLAIVHAVFALKARTVFGVTSVITVVAVFASLPGLIIFTRSQEGLHYTC 178
QY      79 SSHFPYQYQWKNFQTLKI-----LCTRSQKGLHYTC----- 105
Db      179 SSHFPYQYQWKNFQTLKIIVILGLVPLLVVVICYSGLKTLRLCRNEKGRHRAVRLIF 238
QY      106 -----GGSYKGLC-----QEFFGLNCCSSNRLDGHQV 135
Db      239 TIMIVYFLWAPYDIVLLNTFQEFFGLNCCSSNRLDQAMQV 281

Query Match 46.4%; Score 370; DB 6; Length 344;
Best Local Similarity 34.9%; Pred. No. 2.8e-33;
Matches 96; Conservative 9; Mismatches 28; Indels 142; Gaps 7;

QY      1 MDYQVSSPIYDINYTSEPCQKINVKQIAA----- 30

```

```

RA Diop O., Rigoulet J., Barre-Sinoussi F., Fomsgaard A.;
RT "Mutations in CCR5-coding sequences are not associated with HIV
RL carrier status in African nonhuman primates.";
RN AIDS Res. Hum. Retroviruses 15:931-939(1999).
RP [2]
RN SEQUENCE FROM N.A.
RC STRAIN=1410;
RA Mueller-Trutwin M.C., Corbet S., Hansen J., Diop O.,
RA Georges-Courbot M.-C., Barre-Sinoussi F., Fomsgaard A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF035214; AAD44007.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
DR Receptor.
SQ SEQUENCE 352 AA; 40481 MW; 2578A0F2C07A4A65 CRC64;

Query Match 46.3%; Score 369; DB 6; Length 352;
Best Local Similarity 34.3%; Pred. No. 3.7e-33;
Matches 97; Conservative 8; Mismatches 28; Indels 150; Gaps 7;

QY 1 MDYQVSSPIYDINYTSEPCQKINVKQIAA----- 30
DB 1 MDYQVSSPTDYDYTSEPCQKINVKQIAARLLPLYSLVFIQFVGNILVVLINCKR 60
QY 31 -----YKGLC-----AAQWDFGNTMCHQORVGHGHHHSYKCG--- 64
DB 61 LKSMTDIYLLNLAISDLFLLTPVFWAHYAAQWDFGNTM-----LCTRSQKGLHYTC 78
QY 65 ----- 178
DB 119 IILLTDIYLAIVHAFVALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKGLHYTC 178
QY 79 SSHFPYSQYQFKNFOTLKI-----HORVHG----- 105
DB 179 SSHFPYSQYQFKNFOTLKIIVILGLVPLVMVICYSGLKTLRCRNEKXRAVRLIP 238
QY 106 -----GGSYKGLC-----QEFFGLNCCSSNRLDGHQV 135
DB 239 TIMIVVFLFWAPYINVLNLTQEFFGLNCCSSNRLDQAMQV 281

RESULT 6
Q95NC3 PRELIMINARY; PRT: 352 AA.
AC Q95NC3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE C-C chemokine receptor 5.
GN CCR5.
OS Miopithecus talapoin (Talapoin) (Cercopithecus talapoin).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Miopithecus.
OC NCBI_TaxID=36231;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang Y., Ryder O.A., Zhang Y.;
RT "Sequence comparison of the CCR5 gene in primates and primate
RT phylogeny.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF177886; AAK43369.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
DR Receptor.
SQ SEQUENCE 352 AA; 40546 MW; 6464152F3B566AE5 CRC64;

QY 1 MDYQVSSPIYDINYTSEPCQKINVKQIAA----- 30
DB 1 MDYQVSSPTDYDYTSEPCQKINVKQIAARLLPLYSLVFIQFVGNILVVLINCKR 60
QY 31 -----YKGLC-----AAQWDFGNTMCHQORVGHGHHHSYKCG--- 64
DB 61 LKSMTDIYLLNLAISDLFLLTPVFWAHYAAQWDFGNTM-----LCTRSQKGLHYTC 78
QY 65 ----- 178
DB 119 IILLTDIYLAIVHAFVALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKGLHYTC 178
QY 79 SSHFPYSQYQFKNFOTLKI-----HORVHG----- 105
DB 179 SSHFPYSQYQFKNFOTLKIIVILGLVPLVMVICYSGLKTLRCRNEKXRAVRLIP 238
QY 106 -----GGSYKGLC-----QEFFGLNCCSSNRLDGHQV 135
DB 239 TIMIVVFLFWAPYINVLNLTQEFFGLNCCSSNRLDQAMQV 281

RESULT 5
Q9TV50 PRELIMINARY; PRT: 352 AA.
AC Q9TV50;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CC chemokine receptor type 5.
GN CCR5.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
OC NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1410;
RX MEDLINE=99335215; PubMed=10408730;
RA Muller-Trutwin M.C., Corbet S., Hansen J., Georges-Courbot M.C.,

```

```
Query Match 46.0%; Score 367; DB 6; Length 352;
Best Local Similarity 33.9%; Pred. No. 6.2e-33;
Matches 96; Conservative 9; Mismatches 28; Indels 150; Gaps 7;

Db 1 MDYQVSSPIYDINVTSEPCQKINVKQIAA-----30
1 MDYQVSSPIYDINVTSEPCQKINVKQIAARLLPPLYSLVFPGVGNILVILINCKR 60
31 -----YKGLC-----AAAQWDFGNTMCQHQHVHGHHSYKCG---64
61 LKSMTDIYLLNLAIISDLFLLTVPFWAHYAAQWDFGNTMC--RLLTGLYFIFGFSGIF 118
65 -----YKGLC-----LCTRSQKGLHYTC 78
119 IILLTIDRYLAIVHAFKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQKGLHYTC 178
79 SSHFPYSQYQFWKPNFQTLKIVILGLVPLLVWVICYSGLILKTLRCRNEKXKRAVRLIF 238
179 SSHFPYSQYQFWKPNFQTLKIVILGLVPLLVWVICYSGLILKTLRCRNEKXKRAVRLIF 238
106 -----GGSYKGLC-----QEFFGLNCCSSNRLDGHORV 135
239 TIMIVYFLFWAPYINVLNTFQEFFGLNCCSSNRLDQAMQV 281

RESULT 7
C.6771 ID O18771 PRELIMINARY; PRT; 352 AA.
AC O18771;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DE 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DR CC5 receptor (Fragment).
GN CC5.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHCR5-141a;
RA Zhang L., Carruthers C.D., He T., Huang Y., Cao Y., Wang G., Hahn B.,
RT "HIV-1 subtypes, co-receptor usage, and CCR5 polymorphism.";
RL AIDS Res. Hum. Retroviruses 0:0-0(1997).
DR EMBL; AF01539; AAB65739.1;
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 352
SQ SEQUENCE 352 AA; 40466 MW; 3PFAC7ABAE1D4FB CRC64;

Query Match 46.0%; Score 367; DB 6; Length 352;
Best Local Similarity 34.3%; Pred. No. 6.2e-33;
Matches 97; Conservative 8; Mismatches 28; Indels 150; Gaps 7;

Db 1 MDYQVSSPIYDINVTSEPCQKINVKQIAA-----30
1 MDYQVSSPIYDINVTSEPCQKINVKQIAARLLPPLYSLVFPGVGNILVILINCKR 60
31 -----YKGLC-----AAAQWDFGNTMCQHQHVHGHHSYKCG---64
61 LKSMTDIYLLNLAIISDLFLLTVPFWAHYAAQWDFGNTMC--QLLTGLYFIFGFSGIF 118
65 -----YKGLC-----LCTRSQKGLHYTC 78
119 IILLTIDRYLAIVHAFKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQKGLHYTC 178
79 SSHFPYSQYQFWKPNFQTLKIVILGLVPLLVWVICYSGLILKTLRCRNEKXKRAVRLIF 238
106 -----GGSYKGLC-----QEFFGLNCCSSNRLDGHORV 135
239 TIMIVYFLFWAPYINVLNTFQEFFGLNCCSSNRLDQAMQV 281

RESULT 7
C.6771 ID O18771 PRELIMINARY; PRT; 352 AA.
AC O18771;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DE 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DR CC5 receptor (Fragment).
GN CC5.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHCR5-141a;
RA Zhang L., Carruthers C.D., He T., Huang Y., Cao Y., Wang G., Hahn B.,
RT "HIV-1 subtypes, co-receptor usage, and CCR5 polymorphism.";
RL AIDS Res. Hum. Retroviruses 0:0-0(1997).
DR EMBL; AF01539; AAB65739.1;
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 352
SQ SEQUENCE 352 AA; 40466 MW; 3PFAC7ABAE1D4FB CRC64;

Query Match 46.0%; Score 367; DB 6; Length 352;
Best Local Similarity 34.3%; Pred. No. 6.2e-33;
Matches 97; Conservative 8; Mismatches 28; Indels 150; Gaps 7;

Db 1 MDYQVSSPIYDINVTSEPCQKINVKQIAA-----30
1 MDYQVSSPIYDINVTSEPCQKINVKQIAARLLPPLYSLVFPGVGNILVILINCKR 60
31 -----YKGLC-----AAAQWDFGNTMCQHQHVHGHHSYKCG---64
61 LKSMTDIYLLNLAIISDLFLLTVPFWAHYAAQWDFGNTMC--QLLTGLYFIFGFSGIF 118
65 -----YKGLC-----LCTRSQKGLHYTC 78
119 IILLTIDRYLAIVHAFKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQKGLHYTC 178
79 SSHFPYSQYQFWKPNFQTLKIVILGLVPLLVWVICYSGLILKTLRCRNEKXKRAVRLIF 238
106 -----GGSYKGLC-----QEFFGLNCCSSNRLDGHORV 135
239 TIMIVYFLFWAPYINVLNTFQEFFGLNCCSSNRLDQAMQV 281

RESULT 7
C.6771 ID O18771 PRELIMINARY; PRT; 352 AA.
AC O18771;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DE 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DR CC5 receptor (Fragment).
GN CC5.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHCR5-141a;
RA Zhang L., Carruthers C.D., He T., Huang Y., Cao Y., Wang G., Hahn B.,
RT "HIV-1 subtypes, co-receptor usage, and CCR5 polymorphism.";
RL AIDS Res. Hum. Retroviruses 0:0-0(1997).
DR EMBL; AF01539; AAB65739.1;
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 352
SQ SEQUENCE 352 AA; 40466 MW; 3PFAC7ABAE1D4FB CRC64;
```

```
Db 179 SSHFPYSQYQFWKPNFQTLKIVILGLVPLLVWVICYSGLILKTLRCRNEKXKRAVRLIF 238
QY 106 -----GGSYKGLC-----QEFFGLNCCSSNRLDGHORV 135
Db 239 TIMIVYFLFWAPYINVLNTFQEFFGLNCCSSNRLDQAMQV 281

RESULT 8
Q9TSK1 ID Q9TSK1 PRELIMINARY; PRT; 352 AA.
AC Q9TSK1;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DR G-protein coupled chemokine receptor.
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RA Holtkamp N., Baier M., Werner A.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF019379; RAD01639.1;
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 352 AA; 40588 MW; 0F1869D9A6668DBB CRC64;

Query Match 45.9%; Score 366; DB 6; Length 352;
Best Local Similarity 33.9%; Pred. No. 8e-33;
Matches 96; Conservative 9; Mismatches 28; Indels 150; Gaps 7;

QY 1 MDYQVSSPIYDINVTSEPCQKINVKQIAA-----30
Db 1 MDYQVSSPIYDINVTSEPCQKINVKQIAARLLPPLYSLVFPGVGNILVILINCKR 60
QY 31 -----YKGLC-----AAAQWDFGNTMCQHQHVHGHHSYKCG---64
61 LKSMTDIYLLNLAIISDLFLLTVPFWAHYAAQWDFGNTMC--QLLTGLYFIFGFSGIF 118
65 -----YKGLC-----LCTRSQKGLHYTC 78
119 IILLTIDRYLAIVHAFKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQKGLHYTC 178
79 SSHFPYSQYQFWKPNFQTLKIVILGLVPLLVWVICYSGLILKTLRCRNEKXKRAVRLIF 238
106 -----GGSYKGLC-----QEFFGLNCCSSNRLDGHORV 135
239 TIMIVYFLFWAPYINVLNTFQEFFGLNCCSSNRLDQAMQV 281

RESULT 9
Q9TV49 ID Q9TV49 PRELIMINARY; PRT; 352 AA.
AC Q9TV49;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DR CC chemokine receptor type 5 (C-C chemokine receptor 5).
GN CCR5.
OS Cercopithecus galenitus (Agile mangabey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9532;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=4;
RX MEDLINE=99335215; PubMed=10408730;
RA Muller-Trutwin M.C., Corbet S., Hansen J., Georges-Courbot M.C.,
  Diop O., Rigoulet J., Barre-Sinoussi F., Fomsgaard A.;
RT "Mutations in CCR5-coding sequences are not associated with SIV
  carrier status in African nonhuman primates.";
RL AIDS Res. Hum. Retroviruses 15:931-939(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=4;
RA Mueller-Trutwin M.C., Corbet S., Hansen J., Diop O.,
  Georges-Courbot M.-C., Barre-Sinoussi F., Fomsgaard A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Zhang Y., Ryder O.A., Zhang Y.;
RT "Sequence comparison of the CCR5 gene in primates and primate
  phylogeny.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF035215; AAD44008.1; -
DR EMBL; AF177898; AAK43381.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 352 AA; 40479 MW; 5E1504A9B87270B2 CRC64;

Query Match 45.9%; Score 366; DB 6; Length 352;
Best Local Similarity 33.9%; Pred. No. 8e-33;
Matches 96; Conservative 9; Mismatches 28; Indels 150; Gaps 7;

QY 1 MDYQVSSPIYDINVTSEPCOKINVKQIAA----- 30
DB 1 MDYQVSSPTDIDYVTSEPCOKINVKQIAARLLPLYSLVFIFGVGNILVVLINCKR 60
QY 31 -----YKGLG-----AAQDFGNTCMCHQHVHGHHSYKCG-- 64
DB 61 LKSMTDVLLNLAISSDLLFLTLVFWAHYAAQDFGNTMC--OLLTGLYFIFGFSGIFF 118
QY 65 -----LCTRSQKSGLHYTC 78
DB 119 ILLTIDRYLAIVHAFALKARTVTGVTSVITWVAVFASLPGLIIFRSQREGLHYTC 178
QY 79 SSHFPYSQYQWKNFQTLKI-----HORVHCG----- 105
DB 179 SSHFPYSQYQWKNFQTLKIVILGLVPLLVWVICYSGILKTLRCRNEKRRHRAVLIF 238
QY 106 -----GGSYKGLC-----QEFFGLNCGSSNRLDGHORV 135
DB 239 TIMIVYFLFWAPYNIIVLLNTFQEFFGLNCGSSNRLDQAMQV 281

RESULT 10
Q9XTL3 ID Q9XTL3 PRELIMINARY; PRT; 352 AA.
AC Q9XTL3;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE CC chemokine receptor 5.
GN CCR5.
OS Cercopithecus nictitans (white-nosed guenon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Papio.
OX NCBI_TaxID=9555;
RN [1]
RP SEQUENCE FROM N.A.
RA Spearman P.W., Mburu D.N., Graham B.S.;
RT "Differential Utilization of CCR5 Molecules from Three East African
```

```
RT Simian Species by the HIV-1 Envelope Glycoprotein.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF141840; AAD32685.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 352 AA; 40489 MW; 0B47E337C11E2E1E CRC64;

Query Match 45.9%; Score 366; DB 6; Length 352;
Best Local Similarity 33.9%; Pred. No. 8e-33;
Matches 96; Conservative 9; Mismatches 28; Indels 150; Gaps 7;

QY 1 MDYQVSSPIYDINVTSEPCOKINVKQIAA----- 30
DB 1 MDYQVSSPTDIDYVTSEPCOKINVKQIAARLLPLYSLVFIFGVGNILVVLINCKR 60
QY 31 -----YKGLG-----AAQDFGNTCMCHQHVHGHHSYKCG-- 64
DB 61 LKSMTDVLLNLAISSDLLFLTLVFWAHYAAQDFGNTMC--OLLTGLYFIFGFSGIFF 118
QY 65 -----LCTRSQKSGLHYTC 78
DB 119 ILLTIDRYLAIVHAFALKARTVTGVTSVITWVAVFASLPGLIIFRSQREGLHYTC 178
QY 79 SSHFPYSQYQWKNFQTLKI-----HORVHCG----- 105
DB 179 SSHFPYSQYQWKNFQTLKIVILGLVPLLVWVICYSGILKTLRCRNEKRRHRAVLIF 238
QY 106 -----GGSYKGLC-----QEFFGLNCGSSNRLDGHORV 135
DB 239 TIMIVYFLFWAPYNIIVLLNTFQEFFGLNCGSSNRLDQAMQV 281

RESULT 11
Q9TV45 ID Q9TV45 PRELIMINARY; PRT; 352 AA.
AC Q9TV45;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE CC chemokine receptor type 5.
GN CCR5.
OS Cercopithecus nictitans (white-nosed guenon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=36228;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=7;
RX MEDLINE=99335215; PubMed=10408730;
RA Muller-Trutwin M.C., Corbet S., Hansen J., Georges-Courbot M.C.,
  Diop O., Rigoulet J., Barre-Sinoussi F., Fomsgaard A.;
RT "Mutations in CCR5-coding sequences are not associated with SIV
  carrier status in African nonhuman primates.";
RL AIDS Res. Hum. Retroviruses 15:931-939(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=7;
RA Mueller-Trutwin M.C., Corbet S., Hansen J., Diop O.,
  Georges-Courbot M.-C., Barre-Sinoussi F., Fomsgaard A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF035215; AAD44012.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 352 AA; 40658 MW; BC665891ED5661F8 CRC64;
```

```

Query Match      45.9%; Score 366; DB 6; Length 352;
Best Local Similarity 33.9%; Pred. No. 8e-33;
Matches 96; Conservative 9; Mismatches 28; Indels 150; Gaps 7;

QY 1 MDYQVSSPYDINYYTSEPCQKINVKQIAA----- 30
DB 1 MDYQVSSPYDIDYITSEPCQKINVKQIAARLLPLLSLVIFGVGNILVVLINCKR 60
QY 31 -----YKCGLC-----AAQWDFGNTMCHQORVGHGHHHSYKCG--- 64
DB 61 LKSMTDIYLLNLAISDLLFLLTPFWAHYAAQWDFGNTMC--QLLTGLYFIFGFSGIF 118
QY 65 -----LCTRSQKEGLHYTC 78
DB 119 IILLTIDRYLAIVHAFVALKARTVTGVVTSVITWVAVFASLPRIIFTRSQREGLHYTC 178
QY 79 SSHFPYSQYQFWKFNQTLKI-----HORVHG-- 105
DB 179 SSHFPYSQYQFWKFNQTLKIVILGLVPLLVNVCYSGILKTLRCRNEKKRRAVLIF 238
QY 106 -----GGSYKCGLC-----QEFFGLNCCSSNRLDQAMQV 135
DB 239 TIMIVYFLWAFYNYVLLNTFQEFFGLNCCSSNRLDQAMQV 281

RESULT 12
Q95NE8 ID Q95NE8 PRELIMINARY; PRT; 352 AA.
AC Q95NE8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CC chemokine receptor type 5.
GN CCR5.
OS Cercopithecus tantalus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=60712;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99335215; PubMed=10408730;
RA Muller-Trutwin M.C., Corbet S., Hansen J., Georges-Courbot M.C.,
RA Diop O., Rigoulet J., Barre-Sinoussi F., Fomsgaard A.;
RT "Mutations in CCR5-coding sequences are not associated with HIV
RT carrier status in African nonhuman primates";
RL AIDS Res. Hum. Retroviruses 15:931-939(1999).
DR EMBL; AF081577; AAD45495.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
SQ SEQUENCE 352 AA; 40550 MW; 9B078EF04D34DB36 CRC64;

Query Match      45.9%; Score 366; DB 6; Length 352;
Best Local Similarity 33.9%; Pred. No. 8e-33;
Matches 96; Conservative 9; Mismatches 28; Indels 150; Gaps 7;

QY 1 MDYQVSSPYDINYYTSEPCQKINVKQIAA----- 30
DB 1 MDYQVSSPYDIDYITSEPCQKINVKQIAARLLPLLSLVIFGVGNILVVLINCKR 60
QY 31 -----YKCGLC-----AAQWDFGNTMCHQORVGHGHHHSYKCG--- 64
DB 61 LKSMTDIYLLNLAISDLLFLLTPFWAHYAAQWDFGNTMC--QLLTGLYFIFGFSGIF 118
QY 65 -----LCTRSQKEGLHYTC 78
DB 119 IILLTIDRYLAIVHAFVALKARTVTGVVTSVITWVAVFASLPRIIFTRSQREGLHYTC 178
QY 79 SSHFPYSQYQFWKFNQTLKI-----HORVHG-- 105
DB 179 SSHFPYSQYQFWKFNQTLKIVILGLVPLLVNVCYSGILKTLRCRNEKKRRAVLIF 238
QY 106 -----GGSYKCGLC-----QEFFGLNCCSSNRLDQAMQV 135
DB 239 TIMIVYFLWAFYNYVLLNTFQEFFGLNCCSSNRLDQAMQV 281

RESULT 14
Q95NC6 ID Q95NC6 PRELIMINARY; PRT; 352 AA.
AC Q95NC6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CC chemokine receptor 5.
GN CCR5.
OS Trachypithecus johnii (hooded leaf monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```


OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Colobinae;
 OC Trachypithecus.
 OX NCBI_TaxID=66063;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang Y., Ryder O.A., Zhang Y.;
 RT "Sequence comparison of the CCR5 gene in primates and primate
 phylogeny.";
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF177883; AAK43366.1; -;
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm 1; 1.
 DR PRINTS; PR00237; GPCRHOOPS.
 DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
 DR PROSITE; PS00237; G_PROTEIN_RECP_F1_2; 1.
 KW Receptor.
 SQ SEQUENCE 352 AA; 40462 MW; 52824E0322559F7F CRC64;
 Query Match 45.8%; Score 365; DB 6; Length 352;
 Best Local Similarity 33.6%; Pred. No. 1e-32;
 Matches 95; Conservative 10; Mismatches 28; Indels 150; Gaps 7;
 QY 1 MDYQVSSPIYDINYTSEPCQKINVKQIAA----- 30
 DB 1 MDYQVSSPTDYIDYTSQKVNKQIAARLLPPLYSLVFIFGVGNILVVLINCKR 60
 QY 31 -----YKGLC----- 64
 DB 61 LKSMTDIYLLNLALISDLFLLVFPFNAHYAAQWDFGNTMC--QLLTGLYFIFGFSGIF 118
 QY 65 -----LCTRSOKEGLHYTC 78
 DB 119 IILLTIDRYLAIVHAVFALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQREGHYTC 178
 QY 79 SSHFPYSQYQFWKQNFOTLKI-----HQRVHGG----- 105
 DB 179 SSHFPYSQYQFWKQNFOTLKIIVLGLVPLVMVICYSGILKTLRCRNEKKRHRVRLIF 238
 QY 106 -----GGSYKGLC-----OEFGLNCCSSNRLDGHQV 135
 DB 239 TIMIVYFLWAPYINIVLLNTFOEFFGLNCCSSNRLDQAMQV 281

RESULT 15

Q95NC8 PRELIMINARY; PRT; 352 AA.
 AC Q95NC8; 2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE C-C chemokine receptor 5.
 GN CCR5.
 OS Colobus polykomos.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Colobinae;
 OC Colobus.
 OX NCBI_TaxID=9572;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang Y., Ryder O.A., Zhang Y.;
 RT "Sequence comparison of the CCR5 gene in primates and primate
 phylogeny.";
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF177881; AAK43364.1; -;
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm 1; 1.
 DR PRINTS; PR00237; GPCRHOOPS.
 DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
 DR PROSITE; PS00237; G_PROTEIN_RECP_F1_2; 1.
 KW Receptor.
 SQ SEQUENCE 352 AA; 40578 MW; 4366F149C3B4938F CRC64;

Query Match

45.8%; Score 365; DB 6; Length 352;

Best Local Similarity 33.6%; Pred. No. 1e-32;
 Matches 95; Conservative 10; Mismatches 28; Indels 150; Gaps 7;
 QY 1 MDYQVSSPIYDINYTSEPCQKINVKQIAA----- 30
 DB 1 MDYQVSSPTDYIDYTSQKVNKQIAARLLPPLYSLVFIFGVGNILVVLINCKR 60
 QY 31 -----YKGLC----- 64
 DB 61 LKSMTDIYLLNLALISDLFLLVFPFNAHYAAQWDFGNTMC--QLLTGLYFIFGFSGIF 118
 QY 65 -----LCTRSOKEGLHYTC 78
 DB 119 IILLTIDRYLAIVHAVFALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQREGHYTC 178
 QY 79 SSHFPYSQYQFWKQNFOTLKI-----HQRVHGG----- 105
 DB 179 SSHFPYSQYQFWKQNFOTLKIIVLGLVPLVMVICYSGILKTLRCRNEKKRHRVRLIF 238
 QY 106 -----GGSYKGLC-----OEFGLNCCSSNRLDGHQV 135
 DB 239 TIMIVYFLWAPYINIVLLNTFOEFFGLNCCSSNRLDQAMQV 281

Search completed: October 21, 2003, 17:53:38
 Job time : 58.0712 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 17:39:39 ; Search time 57.0712 Seconds
(without alignments)
383.806 Million cell updates/sec

Title: US-10-057-890A-10

Perfect score: 797

Sequence: 1 MDYQVSSPIVDIINYTSEP.....GLNCCSSNRLDGHQVHAA 138

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03:*

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	797	100.0	138	23	ABG32539 Human CCR5-based s
2	797	100.0	157	23	ABG32540 Human CCR5-based s
3	379	47.6	352	18	AAW27407 Human CCR5. Homo
4	379	47.6	352	18	AAW27123 Human Chemokine re
5	379	47.6	352	19	AAW23835 Human CC chemokine
6	379	47.6	352	20	AAW88232 HIV-1 co-receptor
7	379	47.6	352	22	ABBS6342 Non-endogenous hum
8	379	47.6	352	22	AAG80111 Human CCR5 protein
9	379	47.6	352	22	AAW82948 Human HIV-1 co-rec

10	379	47.6	352	22	AAE07039 Human G-protein ch
11	379	47.6	352	22	AAE07048 Human G-protein ch
12	379	47.6	352	22	AAH83354 Human CCR5 protein
13	379	47.6	352	22	AAE04321 Human chemokine re
14	379	47.6	352	22	AAH46858 Human HDGMR10 prot
15	379	47.6	352	22	AAE25811 Human G-protein ch
16	379	47.6	352	23	ABG70597 Human G-protein ch
17	379	47.6	352	23	ABG92883 Human immunoglobul
18	379	47.6	352	23	ABH81054 G-protein chemokine
19	379	47.6	352	23	AAU97152 Human G-protein ch
20	379	47.6	352	23	AAU97152 Human chemokine (C
21	379	47.6	352	23	AAH08343 Human CC chemokine
22	379	47.6	352	23	AAW52829 Human CCR5 Gln 55
23	379	47.6	352	24	ABP97728 Amino acid sequenc
24	379	47.6	352	24	ABG75540 Human G-protein ch
25	379	47.6	352	24	ABF81933 Human C-C chemokine
26	379	47.6	439	20	AAV41280 Fusion protein con
27	374	46.9	371	19	AAW23834 Human CC chemokine
28	373	46.8	352	18	AAW07602 Human G-protein ch
29	373	46.8	352	21	AAV80128 Human G-protein ch
30	373	46.8	352	22	AAE07037 Human G-protein ch
31	373	46.8	352	22	AAE07046 Human G-protein ch
32	373	46.8	352	23	AAE25808 Human G-protein ch
33	373	46.8	352	23	ABG92880 Human G-protein ch
34	373	46.8	352	23	AAU97150 Human G-protein ch
35	363	45.5	352	18	AAW27125 Macaque chemokine
36	359	45.0	352	22	AAG79089 Amino acid sequenc
37	272	34.1	354	19	AAW54037 Mouse CC-CCR5 prot
38	258	32.4	184	18	AAW27406 Inactive human CCR
39	258	32.4	215	18	AAW27408 HIV-1 co-receptor
40	258	32.4	215	20	AAW88238 Human chemokine re
41	183.5	23.0	332	18	AAW26766 HIV-1 co-receptor
42	182.5	22.9	100	20	AAW88231 Chemokine CCR5 ext
43	182	22.8	32	22	AAG80086 Synthetic peptide
44	177	22.2	32	19	AAW43018 Peptide representi
45	177	22.2	32	19	AAW39912

ALIGNMENTS

RESULT 1
ABG32539

ID ABG32539 standard; protein; 138 AA.

XX AC ABG32539;

XX DT 15-NOV-2002 (first entry)

XX DE Human CCR5-based scaffolded fusion protein #1.

XX KW Scaffolded protein; CCR5; HIV; human immunodeficiency virus infection;
KW ECD; extracellular domain; metal chelating motif; zinc finger protein;
KW integral membrane protein; soluble loop; intracellular domain; ICD;
KW gene therapy; immunogen; viral infection; human.

XX OS Homo sapiens.

OS Synthetic.

PV WO200260477-A1.

XX PC 08-AUG-2002.

XX PF 29-JAN-2002; 2002WO-US02377.

XX PR 31-JAN-2001; 2001US-265782P.

XX PR 31-JAN-2001; 2001US-265858P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Coleman TA, Mansfield B;

XX WPI; 2002-643357/69.

```

XX Novel scaffolded fusion polypeptide useful for therapeutic purposes or
PT for screening molecules that bind/activate/inhibit/modulate the
PT polypeptide, comprises a functional polypeptide domain fused to a
PT scaffold domain -
XX
XX Example 1; Page 21; 64pp; English.
XX
XX The invention relates to a scaffolded fusion polypeptide comprising a
CC functional polypeptide domain fused to a scaffold domain, where the
CC functional polypeptide domain corresponds to a soluble loop of an
CC integral membrane protein (e.g. human CCR5, a transmembrane receptor
CC involved in HIV (human immunodeficiency virus) infection).
CC Also included are; (1) a polypeptide comprising a scaffold domain; (2) a
CC nucleic acid encoding the fusion polypeptide; (3) a vector cassette for
CC the expression of the fusion polypeptide comprising an expression region
CC operably linked to a promoter, where the expression region comprises a
CC number of cassettes, each of which encodes a module, domain or strand of
CC the fusion polypeptide and (4) a host cell comprising the vector or
CC nucleic acid. The fusion polypeptide is useful for screening molecules
CC that bind/activate/inhibit/modulate the fusion polypeptide, by expressing
CC the fusion polypeptide from and identifying a molecule that binds to the
CC fusion polypeptide. The fusion polypeptide is useful in diagnostic
CC methods, in assays to identify compounds that interact with loops of
CC fragments of an extracellular domain (ECD) or an intracellular domain
CC (ICD) or to rapidly assay the function of mutated portions of mutant
CC integral membrane proteins without having to produce significant
CC quantities of the entire mutant integral membrane protein, to generate
CC antibodies that recognise the integral membrane proteins from which they
CC are designed, to competitively bind the ligand of a naturally occurring
CC receptor in vitro or in vivo, to display and/or screen soluble domains
CC of protein such as integral membrane proteins, to probe the structure
CC of ECD or ICD, or both, of an integral protein membrane, to modulate the
CC activity of a receptor in vivo, and for treating or preventing viral
CC infection, preferably human HIV infection e.g. by gene therapy using the
CC encoding nucleic acid. The present sequence is a scaffolded protein
CC based on the ECD region of human CCR5 (not defined).
XX
XX Sequence 138 AA;
XX
XX Query Match 100.0%; Score 797; DB 23; Length 138;
XX Best Local Similarity 100.0%; Pred. No. 1.6e-75;
XX Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 MDYQVSSPIVDINYITSEPCQKINVQIAAYKGLCAAQWDFGNTMCOHQRVHGHHS 60
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 1 MDYQVSSPIVDINYITSEPCQKINVQIAAYKGLCAAQWDFGNTMCOHQRVHGHHS 60
XX
XX 61 YKGLCTRSQKEGLHYTCSSHPFYSQYQFWKNPQTLKIHQRVHGGGYSKGLCQEFFGL 120
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 61 YKGLCTRSQKEGLHYTCSSHPFYSQYQFWKNPQTLKIHQRVHGGGYSKGLCQEFFGL 120
XX
XX 121 NNCSSNRLDGHQRVHAA 138
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 121 NNCSSNRLDGHQRVHAA 138
XX
XX RESULT 2
XX ABG32540
XX ID ABG32540 standard; protein; 157 AA.
XX AC ABG32540;
XX
XX 15-NOV-2002 (first entry)
XX
XX Human CCR5-based scaffolded fusion protein #2.
XX
XX Scaffolded protein; CCR5; HIV; human immunodeficiency virus infection;
XX ECD; extracellular domain; metal chelating motif; zinc finger protein;
XX integral membrane protein; soluble loop; intracellular domain; ICD;
XX gene therapy; immunogen; viral infection; human.
XX
XX Homo sapiens.
XX
XX OS

```

```

OS Synthetic.
XX Key Location/Qualifiers
XX Peptide 1..19 /label= Signal_peptide
XX Protein 20..157 /label= Mature_scaffolded_protein
XX
XX WO200260477-A1.
XX
XX 08-AUG-2002.
XX
XX 29-JAN-2002; 2002WO-US02377.
XX
XX 31-JAN-2001; 2001US-265782P.
XX 31-JAN-2001; 2001US-265858P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Coleman TA, Mansfield B;
XX WPI; 2002-643357/69.
XX
XX Novel scaffolded fusion polypeptide useful for therapeutic purposes or
XX for screening molecules that bind/activate/inhibit/modulate the
XX polypeptide, comprises a functional polypeptide domain fused to a
XX scaffold domain -
XX
XX Example 2; Page 41; 64pp; English.
XX
XX The invention relates to a scaffolded fusion polypeptide comprising a
XX functional polypeptide domain fused to a scaffold domain, where the
XX functional polypeptide domain corresponds to a soluble loop of an
XX integral membrane protein (e.g. human CCR5, a transmembrane receptor
XX involved in HIV (human immunodeficiency virus) infection).
XX Also included are; (1) a polypeptide comprising a scaffold domain; (2) a
XX nucleic acid encoding the fusion polypeptide; (3) a vector cassette for
XX the expression of the fusion polypeptide comprising an expression region
XX operably linked to a promoter, where the expression region comprises a
XX number of cassettes, each of which encodes a module, domain or strand of
XX the fusion polypeptide and (4) a host cell comprising the vector or
XX nucleic acid. The fusion polypeptide is useful for screening molecules
XX that bind/activate/inhibit/modulate the fusion polypeptide, by expressing
XX the fusion polypeptide from and identifying a molecule that binds to the
XX fusion polypeptide. The fusion polypeptide is useful in diagnostic
XX methods, in assays to identify compounds that interact with loops of
XX fragments of an extracellular domain (ECD) or an intracellular domain
XX (ICD) or to rapidly assay the function of mutated portions of mutant
XX integral membrane proteins without having to produce significant
XX quantities of the entire mutant integral membrane protein, to generate
XX antibodies that recognise the integral membrane proteins from which they
XX are designed, to competitively bind the ligand of a naturally occurring
XX receptor in vitro or in vivo, to display and/or screen soluble domains
XX of protein such as integral membrane proteins, to probe the structure
XX of ECD or ICD, or both, of an integral protein membrane, to modulate the
XX activity of a receptor in vivo, and for treating or preventing viral
XX infection, preferably human HIV infection e.g. by gene therapy using the
XX encoding nucleic acid. The present sequence is a scaffolded protein
XX based on the ECD region of human CCR5 (not defined).
XX
XX Sequence 157 AA;
XX
XX Query Match 100.0%; Score 797; DB 23; Length 157;
XX Best Local Similarity 100.0%; Pred. No. 1.8e-75;
XX Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 MDYQVSSPIVDINYITSEPCQKINVQIAAYKGLCAAQWDFGNTMCOHQRVHGHHS 60
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 20 MDYQVSSPIVDINYITSEPCQKINVQIAAYKGLCAAQWDFGNTMCOHQRVHGHHS 79
XX
XX 61 YKGLCTRSQKEGLHYTCSSHPFYSQYQFWKNPQTLKIHQRVHGGGYSKGLCQEFFGL 120
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 80 YKGLCTRSQKEGLHYTCSSHPFYSQYQFWKNPQTLKIHQRVHGGGYSKGLCQEFFGL 139
XX

```

```

Qy 121 NCSSNRLDGHORVHAA 138
Db 140 NCSSNRLDGHORVHAA 157

RESULT 3
AAW27407
ID AAW27407 standard; Protein; 352 AA.
XX AC
XX AC
XX 14-APR-1998 (first entry)
DT XX
DE Human CCR5.
XX
XX Human Cys-Cys chemokine receptor 5; CCR5;
KW human immunodeficiency virus; type 1; type 2; HIV-1; HIV-2;
KW diagnosis; treatment; prevention;
KW inflammatory disease; rheumatoid arthritis; glomerulonephritis;
KW asthma; idiopathic pulmonary fibrosis; psoriasis; viral infection;
KW cancer; atherosclerosis; autoimmune disorder.
XX
OS Homo sapiens.
XX
XX WO9732019-A2.
XX PD 04-SEP-1997.
XX
XX 28-FEB-1997; 97WO-BE00023.
XX PF
XX PR 06-AUG-1996; 96EP-0870102.
XX PR 01-MAR-1996; 96EP-0870021.
XX
XX (EURO-) EUROSREEN SA.
XX PA
XX
XX Libert F, Parmentier M, Samson M, Vassart G;
XX WPI; 1997-479829/44.
XX DR N-PSDB; AAT90117.
XX
XX Active and inactive forms of human CC chemokine receptor CCR-5 -
XX useful to diagnose, prevent and/or treat inflammatory disorders,
XX autoimmune disease and viral infection
XX
XX Claim 4; Fig 1b-c; 94pp; English.
XX
XX The present sequence is human CC (Cys-Cys) chemokine receptor
XX 5 (CCR5), which is stimulated by MIP-1 alpha, MIP-1 beta or RANTES
XX chemokines, but not by monocyte chemoattractant protein 1 (MCP-1),
XX MCP-2, MCP-3, interleukin-8 (IL-8) or growth related gene product
XX alpha (GRO alpha) chemokines. Active CCR-5 is also a receptor of
XX human immunodeficiency virus type 1 or type 2 (HIV-1 or HIV-2).
XX CCR5 or its cDNA can be used to diagnose, treat and/or prevent
XX inflammatory diseases, e.g. rheumatoid arthritis,
XX glomerulonephritis, asthma, idiopathic pulmonary fibrosis and
XX psoriasis, viral infections, especially HIV-1 or HIV-2 infection,
XX cancer, atherosclerosis and autoimmune disorders.
XX
SQ Sequence 352 AA;

Query Match 47.6%; Score 379; DB 18; Length 352;
Best Local Similarity 35.0%; Pred. No. 2.4e-31;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

Qy 1 MDYQVSSPIYDINVTSEPCOKINVKQIAA----- 30
Db 1 MDYQVSSPIYDINVTSEPCOKINVKQIAARLLPLYSLVIFFGVGNMLVILINCKR 60
Qy 31 -----YKCGLC-----AAAQWDFGNTMCQHORVHGHHHSYKCG--- 64
Db 61 LKSMTDIYLLNLATSLDFLLTVPFWAHYAAQWDFGNTMC--QLLTGLYFIFGPFSGIFF 118

Qy 65 -----LCTRSQKGLHYTC 78
Db 119 IILLTDRLAVVHVPALKARTVTFGVTSVITWVAVFASLPGIIFTRSQKGLHYTC 178
Qy 79 SSHFPYSQYQFKNFQTLKI-----HQRVHGG----- 105
Db 179 SSHFPYSQYQFKNFQTLKIVILGLVPLLVVVICYSGILKTLRLCRNEKRRHRAVRLIF 238
Qy 106 -----CGSYKGLC-----OFFGLNCCSSNRLDGHORV 135
Db 239 TIMIVYELFWAPYNIVLLNTFOEFFGLNCCSSNRLDQAMQV 281

RESULT 4
AAW27123
ID AAW27123 standard; Protein; 352 AA.
XX AC
XX AAW27123;
XX 14-DEC-1997 (first entry)
DT XX
DE Human chemokine receptor 88C.
XX
XX Chemokine receptor 88C; atherosclerosis; rheumatoid arthritis;
KW tumour; asthma; viral infection; AIDS; inflammation;
KW autoimmune disease; therapy; diagnosis; leukocyte trafficking;
KW G protein coupled receptor; ligand; modulator; antibody; human.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT Domain 1..32 /label= Extracellular_domain
XX FT Domain 56..67 /label= Intracellular_domain
XX FT Domain 89..112 /label= Extracellular_domain
XX FT Domain 125..145 /label= Intracellular_domain
XX FT Domain 166..191 /label= Extracellular_domain
XX FT Domain 213..235 /label= Intracellular_domain
XX FT Domain 259..280 /label= Extracellular_domain
XX FT Domain 301..352 /label= Intracellular_domain
XX
XX WO9722698-A2.
XX PN
XX 26-JUN-1997.
XX PD
XX 20-DEC-1996; 96WO-US20759.
XX PF
XX 07-JUN-1996; 96US-0661393.
XX PR 20-DEC-1995; 95US-0575967.
XX
XX (ICOS-) ICOS CORP.
XX PA
XX Gray PW, Raport CJ, Schweickart VL;
XX
XX WPI; 1997-341689/31.
XX DR N-PSDB; AAT85161.
XX
XX New nucleic acid encoding chemokine receptors 88-2B and 88C - used
XX to modulate leukocyte trafficking, e.g. for treatment of
XX inflammation, tumours, viral infections, autoimmune diseases, etc.
XX
XX Claim 16; Page 47-48; 65pp; English.
XX
XX This polypeptide sequence comprises novel human chemokine receptor
XX 88C, a G protein coupled receptor that is involved in leukocyte
XX trafficking. Its amino sequence was deduced from a cDNA clone

```

CC (AA785161) isolated from a macrophage library. It shows 62% identity
 CC to CCR5. Chemokine receptor 88-2B (see AAW2124) has also been
 CC identified. 88C and 88-2B receptors and their polypeptide fragments
 CC can be produced in transformed host cells. The receptors, peptides
 CC comprising one or more of the extracellular or intracellular
 CC domains, and anti-receptor antibodies can be used to modulate
 CC receptor activities, particularly ligand and G protein binding, and
 CC are potentially useful in the treatment of
 CC atherosclerosis, rheumatoid arthritis, tumours, asthma, viral
 CC infection, AIDS, inflammatory conditions, pathological immune
 CC response, abnormal haematopoietic processes etc.

XX Sequence 352 AA;

Query Match 47.6%; Score 379; DB 18; Length 352;

Best Local Similarity 35.0%; Pred. No. 2.4e-31;

Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

QY 1 MDYQVSSPIYDINYTSEPCQKINVKQIAA----- 30
 DB 1 MDYQVSSPIYDINYTSEPCQKINVKQIAARLLPPLYSLVFIFGVGNMLVILLINCKR 60
 QY 31 -----YKCGLC-----AAAQWDFGNTMCOHQRVHGHHSYKCG--- 64
 DB 61 LKSMTDIYLLNLAISDLFLLTPFWAHYAAQWDFGNTMC--QLLTGLYFIFGFSGIF 118
 QY 65 -----YKCGLC-----LCTRSQEGHLYTC 78
 DB 119 ILLTIDRYLAVVHAFKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGHLYTC 178
 QY 79 SSHFPYSQYQFKNFOTLKI-----HORVHGG----- 105
 DB 179 SSHFPYSQYQFKNFOTLKIIVGLVPLVMVICYGILKTLRCRNEKRAVRLLIF 238
 QY 106 -----GSGYKCGLC-----QEFFGLNCSNRLDGHQRV 135
 DB 239 TIMIVYFLWAPYNIIVLLNTTQEFFGLNCSNRLDQAMQV 281

RESULT 5

AAW23835

ID AAW23835 standard; Protein; 352 AA.

AC AAW23835;

DT 08-JUN-1998 (first entry)

XX Human CC chemokine receptor 5 (CCR5).

DE CC chemokine receptor 5; CCR5; G-protein coupled receptor;

KW human immunodeficiency virus; HIV; CD4; AIDS; therapy;

XX transgenic animal.

OS Homo sapiens.

XX Key Location/Qualifiers

FH Domain 29..55

FT /label= I

FT /note= "transmembrane domain"

FT Region 109..120

FT /note= "extracellular loop-1 (Claim 19)"

FT Domain 104..126

FT /label= III

FT /note= "transmembrane domain"

FT Domain 143..171

FT /label= IV

FT /note= "transmembrane domain"

FT Region 187..210

FT /note= "extracellular loop-2 (Claim 19)"

FT Domain 194..219

FT /label= V

FT /note= "transmembrane domain"

FT Domain 238..258

FT /label= VI
 FT /note= "transmembrane domain"
 FT Region 261..276
 FT /note= "extracellular loop-3 (Claim 19)"
 FT Domain 277..300
 FT /label= VII
 FT /note= "transmembrane domain"

XX W09745543-A2.

PN 04-DEC-1997.

XX 28-MAY-1997; 97MO-US09586.

XX 28-MAY-1996; 96US-0018508.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Alkharib G, Berger EA, Broder CC, Combadiere C;

XX Feng Y, Kennedy PE, Murphy PM;

XX WPI; 1998-032650/03.

XX N-PSDB; AAT76920.

CC chemokine receptor 5 polypeptide - used to inhibit membrane
 fusion between HIV and a target cell

XX Claim 68; Fig 1C; 70pp; English.

XX This protein sequence comprises of a novel human macrophage-selective
 CC chemokine receptor that has been designated CCR5. The sequence
 CC was deduced from an isolated cDNA clone (see AAT76920). An Alu27Leu
 CC variant (see W38340 of CCR5) was also identified. The susceptibility
 CC of human macrophages to HIV infection depends on cell surface
 CC expression of CD4 and CCR5. CCR5 is a member of the 7-transmembrane
 CC superfamily of G-protein coupled cell surface molecules. It plays
 CC an essential role in the membrane fusion step of infection by some
 CC HIV isolates. The establishment of stable, non-human cell lines
 CC and transgenic mammals having cells that coexpress human CD4 and
 CC CCR5 provides valuable tools for research of HIV infection.

CC Antibodies that bind to CCR5, CCR5 variants, and CCR5-binding
 CC agents capable of blocking membrane fusion between HIV and target
 CC cells represent potential anti-HIV therapeutics for macrophage
 CC tropic strains of HIV.

XX Sequence 352 AA;

Query Match 47.6%; Score 379; DB 19; Length 352;

Best Local Similarity 35.0%; Pred. No. 2.4e-31;

Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

QY 1 MDYQVSSPIYDINYTSEPCQKINVKQIAA----- 30

DB 1 MDYQVSSPIYDINYTSEPCQKINVKQIAARLLPPLYSLVFIFGVGNMLVILLINCKR 60

QY 31 -----YKCGLC-----AAAQWDFGNTMCOHQRVHGHHSYKCG--- 64

DB 61 LKSMTDIYLLNLAISDLFLLTPFWAHYAAQWDFGNTMC--QLLTGLYFIFGFSGIF 118

QY 65 -----YKCGLC-----LCTRSQEGHLYTC 78

DB 119 ILLTIDRYLAVVHAFKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGHLYTC 178

QY 79 SSHFPYSQYQFKNFOTLKI-----HORVHGG----- 105

DB 179 SSHFPYSQYQFKNFOTLKIIVGLVPLVMVICYGILKTLRCRNEKRAVRLLIF 238

QY 106 -----GSGYKCGLC-----QEFFGLNCSNRLDGHQRV 135

DB 239 TIMIVYFLWAPYNIIVLLNTTQEFFGLNCSNRLDQAMQV 281

RESULT 6

CC endocrine, motor or urogenital systems or skin are affected, and bone
 CC marrow diseases. The products of the invention are chemokine derivatives
 CC which have cytostatic, antiinflammatory, antiasthmatic,
 CC immunosuppressive, dermatological, antirheumatic, antiarthritic.
 CC Chemokines act on specific tumor and inflammatory cells through a
 CC constellation of chemokine receptors (CR), which control migration and
 CC proliferation of these cells. AAG80045-AAG80128 represent human chemokine
 CC fragments used to illustrate the method of the invention.
 XX Sequence 352 AA;
 SQ
 Query Match 47.6%; Score 379; DB 22; Length 352;
 Best Local Similarity 35.0%; Pred. No. 2.4e-31;
 Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;
 QY 1 MDYQVSSPIVDINYYTSEPCQKINVKQIAA----- 30
 DB 1 MDYQVSSPIVDINYYTSEPCQKINVKQIAA----- 60
 QY 31 -----YKGLC----- 64
 DB 61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAQWDFGNTMC--QLLTGLYFTGFFSGIFF 118
 QY 65 -----LCTRSQKEGLHYTC 178
 DB 119 IILLTDRIYLAHVAVFALKARTVTFGVVTSVITVVAVFASLPGLIIFTRSQKEGLHYTC 178
 QY 79 SSHFPYSQYQFWMKQFQTLKI-----HQRVHGG----- 105
 DB 179 SSHFPYSQYQFWMKQFQTLKI-----HQRVHGG----- 238
 QY 106 -----GGSYKGLC-----OEFGLNCCSSNRLDGHQV 135
 DB 239 TIMIVYFLFWAPYNIIVLLNTFOEFGFLNCCSSNRLDQAMQV 281
 RESULT 9
 AAB82948
 ID AAB82948 standard; Protein; 352 AA.
 XX
 AC AAB82948;
 XX
 DT 21-DEC-2001 (first entry)
 XX Human HIV-1 co-receptor CCR5.
 DE CCR5; chemokine; co-receptor; human immunodeficiency virus type 1;
 KW HIV-1; infection; therapy; vaccine; anti-HIV-1.
 KW Homo sapiens.
 OS
 PH Key Location/Qualifiers
 FT Binding-site 2..18
 FT /note= "binds to HIV-1 gp120"
 XX
 PN WO200164710-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 28-FEB-2001; 2001WO-US06699.
 XX
 PR 29-FEB-2000; 2000US-185667P.
 PR 19-MAY-2000; 2000US-205839P.
 PR 07-FEB-2001; 2001US-267231P.
 XX
 PA (PROG-) PROGENICS PHARM INC.
 PA (AARO-) AARON DIAMOND AIDS RES CENT.
 XX
 PI Dragic T, Olson WC;
 XX
 DR WPI; 2001-611273/70.
 DR N-PSDB; AAH26903.
 XX

Query Match 47.6%; Score 379; DB 22; Length 352;
 Best Local Similarity 35.0%; Pred. No. 2.4e-31;
 Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;
 QY 1 MDYQVSSPIVDINYYTSEPCQKINVKQIAA----- 30
 DB 1 MDYQVSSPIVDINYYTSEPCQKINVKQIAA----- 60
 QY 31 -----YKGLC----- 64
 DB 61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAQWDFGNTMC--QLLTGLYFTGFFSGIFF 118
 QY 65 -----LCTRSQKEGLHYTC 178
 DB 119 IILLTDRIYLAHVAVFALKARTVTFGVVTSVITVVAVFASLPGLIIFTRSQKEGLHYTC 178
 QY 79 SSHFPYSQYQFWMKQFQTLKI-----HQRVHGG----- 105
 DB 179 SSHFPYSQYQFWMKQFQTLKI-----HQRVHGG----- 238
 QY 106 -----GGSYKGLC-----OEFGLNCCSSNRLDGHQV 135
 DB 239 TIMIVYFLFWAPYNIIVLLNTFOEFGFLNCCSSNRLDQAMQV 281
 RESULT 8
 AAG80111
 ID AAG80111 standard; Protein; 352 AA.
 XX
 AC AAG80111;
 XX
 DT 17-JAN-2002 (first entry)
 XX Human CCR5 protein.
 DE
 XX
 KW Chemokine; tumour diagnosis; colorectal; prostatic; organ rejection;
 KW inflammation; autoimmune disease; metastasis; bronchial asthma; lupus;
 KW chronic bowel inflammation; rheumatoid arthritis; cytostatic;
 KW antiinflammatory; antiasthmatic; immunosuppressive; dermatological;
 KW antirheumatic; antiarthritic.
 XX
 OS Homo sapiens.
 OS
 PN WO200172830-A2.
 XX
 PD 04-OCT-2001.
 XX
 PF 02-APR-2001; 2001WO-EP03708.
 XX
 PR 31-MAR-2000; 2000DE-1016013.
 XX
 PA (IPFP-) IPF PHARM GMBH.
 PA (FORS/) FORSMANN U.
 XX
 PI Forsemann W, Adermann K, Heitland A, Spodsberg N;
 XX
 DR WPI; 2001-626256/72.
 XX
 PT Diagnostic agent containing two or more receptor-specific ligands,
 PT useful for detecting tumors, inflammation etc., also therapeutic use of
 PT ligand inhibitors -
 XX
 PS Disclosure; Page 10; 26pp; German.
 XX
 CC This invention describes a novel diagnostic agent (A) comprising at least
 CC two different ligands (I) for receptors (II) that are implicated in
 CC disease. (A) are used for the diagnosis of tumors (especially colorectal
 CC or prostatic), organ rejection, inflammation and autoimmune diseases.
 CC Also inhibitors of (I) are used therapeutically against tumors (and their
 CC metastases), inflammation (particularly bronchial asthma or chronic bowel
 CC inflammation), or autoimmune diseases (rheumatoid arthritis or lupus),
 CC where the (cardio)vascular, lymphatic, respiratory, nervous, digestive,

CC binding. The method is used to identify an agent capable of modulating
 CC the interaction of CCR5 and gp120. An agent identified by the method
 CC is used to prepare a pharmaceutical composition for the treatment of a
 CC disease or condition associated with CCR5 and gp120 interaction, to treat
 CC a subject with a disease or condition associated with CCR5 and gp120
 CC interaction, and for preparing a pharmaceutical for treating human
 CC immunodeficiency virus (HIV). It can also be used to treat anti-inflammatory
 CC diseases. The method is commercially useful, amenable to high throughput
 CC screening, and detects interaction of gp120 with cells expressing only
 CC CCR5.

XX Sequence 352 AA;
 SQ Query Match 47.6%; Score 379; DB 22; Length 352;
 Best Local Similarity 35.0%; Pred. No. 2.4e-31;
 Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;
 QY 1 MDYQVSSPIYDINYTSEPCOKINVKQIAA----- 30
 DB 1 MDYQVSSPIYDINYTSEPCOKINVKQIAAARLLPPLYSLVFIFFGVGNMLVILINCKR 60
 QY 31 -----YKCGLC----- 64
 DB 61 LKSMTDIYLLNLAIISDLFLLTVFPAHYAAQAQDFGNTWC--QLLTGLYFIFFSGIFF 118
 QY 65 -----LCTRSQKGLHYTC 78
 DB 119 IILLTIDRYLAVHAFALKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQKGLHYTC 178
 QY 79 SSHFPYSQYQFWMKFNFTLKI----- 105
 DB 179 SSHFPYSQYQFWMKFNFTLKIIVILGLVPLLVVVICYSGLIKTLRCRNEKKRRAVRLIF 238
 QY 106 -----GSGYKCGLC----- 135
 DB 239 TIMIVYFLWAPYNIVLLNTFQEFFGLNCGSSNRLDQAMQV 281

RESULT 13
 AA04321
 ID AAE04321 standard; Protein; 352 AA.
 AC AAE04321;
 XX 04-SEP-2001 (first entry)
 XX Human chemokine receptor (CCR), CC-CR-5 related protein #2.
 XX Human; transformed mammalian cell; CD4; reporter gene; translocation;
 KW human immuno deficiency virus; HIV; long terminal repeat; LTR; therapy;
 KW chemokine receptor; CCR; cellular dysfunction; HIV infection; cofactor;
 KW CC-CR-5; envelope glycoprotein; anti-HIV.
 XX Homo sapiens.
 XX US6258527-B1.
 XX 10-JUL-2001.
 XX 21-MAY-1997; 97US-0861105.
 XX 20-MAY-1996; 96US-0017157.
 PR 19-JUN-1996; 96US-0020043.
 PR 19-MAY-1997; 97US-0858660.
 XX (AARO-) AARON DIAMOND AIDS RES CENT.
 PA (UTNY) UNIV NEW YORK STATE.
 XX Littman DR, Deng H, Ellmeier W, Landau NR, Liu R;
 DR WPI; 2001-417127/44.
 DR N-PSDB; AAD08577.

PT Transformed mammalian cell (1) that contains a CD4 gene, reporter gene
 PT and HIV LTR for identification of drugs and antibodies for treatment of
 HIV -
 XX Disclosure; Column 47-50; 37pp; English.
 XX The present invention relates to a transformed mammalian cell that
 CC contains a gene encoding CD4, a construct encoding a reporter gene under
 CC the regulation of an human immuno deficiency virus (HIV) long terminal
 CC repeat (LTR) and that has been transduced with a vector encoding a
 CC human chemokine receptor (CCR) where the CD4 and the CCR are present on
 CC the cell surface of transformed mammalian cell. The invention is useful
 CC for identifying drugs or antibodies that interfere with the
 CC translocation of HIV into transformed mammalian cell or for identifying
 CC a human chemokine receptor that facilitates the infection of a
 CC particular HIV strain into the transformed mammalian cell. Compounds
 CC identified can be used to treat cellular dysfunction and to prevent or
 CC combat HIV infection. The present sequence is a human chemokine receptor
 CC (CCR), CC-CR-5 related protein. CC-CR-5 is the principal cofactor for
 CC entry mediated by the envelope glycoproteins of primary macrophage-tropic
 CC strains of HIV-1.

XX Sequence 352 AA;
 SQ Query Match 47.6%; Score 379; DB 22; Length 352;
 Best Local Similarity 35.0%; Pred. No. 2.4e-31;
 Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;
 QY 1 MDYQVSSPIYDINYTSEPCOKINVKQIAA----- 30
 DB 1 MDYQVSSPIYDINYTSEPCOKINVKQIAAARLLPPLYSLVFIFFGVGNMLVILINCKR 60
 QY 31 -----YKCGLC----- 64
 DB 61 LKSMTDIYLLNLAIISDLFLLTVFPAHYAAQAQDFGNTWC--QLLTGLYFIFFSGIFF 118
 QY 65 -----LCTRSQKGLHYTC 78
 DB 119 IILLTIDRYLAVHAFALKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQKGLHYTC 178
 QY 79 SSHFPYSQYQFWMKFNFTLKI----- 105
 DB 179 SSHFPYSQYQFWMKFNFTLKIIVILGLVPLLVVVICYSGLIKTLRCRNEKKRRAVRLIF 238
 QY 106 -----GSGYKCGLC----- 135
 DB 239 TIMIVYFLWAPYNIVLLNTFQEFFGLNCGSSNRLDQAMQV 281

RESULT 14
 AAB46858
 ID AAB46858 standard; Protein; 352 AA.
 AC AAB46858;
 XX 16-AUG-2001 (updated)
 DT 02-AUG-2001 (updated)
 DT 04-MAY-2001 (first entry)
 XX Human HDGMR10 protein.
 XX HDGMR10; human; G-protein chemokine receptor; antiinflammatory;
 KW immunomodulatory; anticoagulant; antiallergic; immunosuppressive;
 KW cytostatic; antiparasitic; antipruritic; antirheumatic; antiathritic;
 KW vasotropic; gene therapy; haematopoiesis; wound healing; coagulation;
 KW angiogenesis; solid tumour; infection; leukemia; growth factor activity;
 KW T-cell mediated autoimmune disease; psoriasis; allergy; atherosclerosis;
 KW anaphylaxis; malignancy; inflammation; histamine; IgE; sialic acid; shock;
 KW immunoglobulin E-mediated allergic reaction; rheumatoid arthritis;
 KW prostaglandin-independent fever; bone marrow failure; sarcoidosis;
 KW hyper-eosinophilic syndrome; vulnery.
 XX Homo sapiens.

```
XX US2001000241-A1.
XX 12-APR-2001.
XX 29-NOV-2000; 2000US-0725285.
XX 06-JUN-1995; 95US-0466343.
XX 18-NOV-1998; 98US-0195662.
XX 25-JUN-1999; 99US-0339912.
XX (LIYY/) LI Y.
XX (RUBE/) RUBEN S M.
XX Li Y, Ruben SM;
XX WPI; 2001-226317/23.
XX N-PSDB; AAF26390.
XX New human G-protein chemokine receptor polypeptides and
XX polynucleotides, useful for identifying (ant)agonists to the G-protein
XX chemokine receptor -
XX Claim 1a; Page 15; 22pp; English.
XX This invention describes a novel receptor polypeptide (I) selected from
XX (i) a fully defined 329 amino acid sequence (II) fully disclosed in the
XX specification; and (ii) a polypeptide encoded by the cDNA contained in a
XX plasmid, and fragments, analogs and derivatives of the polypeptide. The
XX products of the invention have antiinflammatory, immunomodulatory,
XX anticoagulant, antiallergic, immunosuppressive, vulnerary, cytostatic,
XX antiparasitic, antipsoriatic, antirheumatic, antiarthritic and vasotropic
XX activity and can be used for gene therapy. The G-protein chemokine
XX receptors, HDGNR10, (I) are useful for screening for compounds which
XX activate or inhibit activation of (I). The products of the invention can
XX also be used for stimulating haematopoiesis, wound healing, coagulation,
XX angiogenesis, treating solid tumours, chronic infections, leukemia, and
XX T-cell mediated autoimmune diseases, parasitic infections, psoriasis, and
XX stimulating growth factor activity. HDGNR10 is useful for treating
XX allergy, atherosclerosis, anaphylaxis, malignancy, chronic and acute
XX inflammation, histamine and immunoglobulin E (IgE)-mediated allergic
XX reactions, prostaglandin-independent fever, bone marrow failure,
XX silicosis, sarcoidosis, rheumatoid arthritis, shock and
XX hyper-eosinophilic syndrome.
XX (N.B. This record was resubmitted to correct errors in the keyword
XX formatting).
XX Sequence 352 AA;
XX
XX Query Match 47.6%; Score 379; DB 22; Length 352;
XX Best Local Similarity 35.0%; Pred. No. 2.4e-31;
XX Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;
XX
XX QY 1 MDYQVSSPIYDINTYSEPCQKINVKQIAA----- 30
XX 1 MDYQVSSPIYDINTYSEPCQKINVKQIAARLLPPLYSLVFIFGFGVGNMLVILINCKR 60
XX
XX QY 31 -----YKGLC----- 64
XX 61 LKSMTDIYLLNLAI S D L F L L T V P F W A H Y A A Q W D F G N T W C - Q L L T G L Y F I G F S G I F F 118
XX 65 -----LCTRSQKEGLHYTC 78
XX 119 I I L L T D R Y L A V H A V F A L K A R T V T F G W T S V I T W V A V P A S L P G I F T R S Q K E G L H Y T C 178
XX 79 S S H F P Y S Q Y F W K N F O T L K I - - - - - H Q R V H G - - - - - 105
XX 179 S S H F P Y S Q Y F W K N F O T L K I V I L G L V L P L V M V I C Y S G I L K T L L R C N E K Q R H A V R L I F 238
XX 106 -----GGSYKCGLC-----O E F F L A N C S S N R L D G H Q R V 135
XX 239 T I M V Y F L F W A P Y N I V L L N T F O E F F L A N C S S N R L D Q M Q V 281
XX
XX RESULT 15
XX AAE25811
XX ID AAE25811 standard; Protein; 352 AA.
XX AC AAE25811;
XX DT 24-FEB-2003 (first entry)
XX DE Human G-protein chemokine receptor (CCR5), HDGNR10 #2.
XX KW Human; G-protein chemokine receptor; CCR5; HDGNR10 protein; cancer;
XX inflammation; viral infection; autoimmune disease; neurodegeneration;
XX rheumatoid arthritis; Pneumocystis carinii infection; Kaposi's sarcoma;
XX hyperproliferative disease; receptor.
XX OS Homo sapiens.
XX PN US2002061834-A1.
XX PD 23-MAY-2002.
XX PF 09-FEB-2001; 2001US-0779880.
XX PR 09-FEB-2000; 2000US-181258P.
XX PR 09-MAR-2000; 2000US-187999P.
XX PR 22-SEP-2000; 2000US-234336P.
XX PA (ROSE/) ROSEN C A.
XX PA (ROSC/) ROSCHKE V.
XX PA (LIYY/) LI Y.
XX PA (RUBE/) RUBEN S M.
XX PI Rosen CA, Roschke V, Li Y, Ruben SM;
XX WPI; 2002-499674/53.
XX DR N-PSDB; AAD42426.
XX
XX New nucleic acid encoding antibodies to the human CCR5 receptor
XX HDGNR10, useful for treatment, prevention and diagnosis of e.g. cancer,
XX also related antibodies -
XX Disclosure; Page 170; 186pp; English.
XX The invention relates to human G-protein chemokine receptor (CCR5),
XX HDGNR10 proteins and nucleic acid molecules encoding such proteins.
XX CCR5 antibodies are used for the treatment or prevention of inflammation,
XX defective or aberrant chemotaxis of immune cells or T cell antigen-
XX presenting cell interaction, viral infections (specifically human immune
XX deficiency (including its early stages), cytomegalo or pox viruses),
XX autoimmune disease, rheumatoid arthritis, neurodegeneration, Pneumocystis
XX carinii infection, Kaposi's sarcoma or any condition associated with
XX aberrant expression of CCR5 or their ligands. They are also used for the
XX detection, diagnosis, prognosis and monitoring of cancers or other
XX hyperproliferative diseases. The present sequence is human G-protein
XX chemokine receptor (CCR5), HDGNR10 DNA.
XX
XX SQ Sequence 352 AA;
XX
XX Query Match 47.6%; Score 379; DB 23; Length 352;
XX Best Local Similarity 35.0%; Pred. No. 2.4e-31;
XX Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;
XX
XX QY 1 MDYQVSSPIYDINTYSEPCQKINVKQIAA----- 30
XX 1 MDYQVSSPIYDINTYSEPCQKINVKQIAARLLPPLYSLVFIFGFGVGNMLVILINCKR 60
XX
XX QY 31 -----YKGLC----- 64
XX 61 LKSMTDIYLLNLAI S D L F L L T V P F W A H Y A A Q W D F G N T W C - Q L L T G L Y F I G F S G I F F 118
XX 65 -----LCTRSQKEGLHYTC 78
XX 119 I I L L T D R Y L A V H A V F A L K A R T V T F G W T S V I T W V A V P A S L P G I F T R S Q K E G L H Y T C 178
XX 79 S S H F P Y S Q Y F W K N F O T L K I - - - - - H Q R V H G - - - - - 105
XX 179 S S H F P Y S Q Y F W K N F O T L K I V I L G L V L P L V M V I C Y S G I L K T L L R C N E K Q R H A V R L I F 238
XX 106 -----GGSYKCGLC-----O E F F L A N C S S N R L D G H Q R V 135
XX 239 T I M V Y F L F W A P Y N I V L L N T F O E F F L A N C S S N R L D Q M Q V 281
XX
```

```

Db      . 119 IILLTDRYLAVHVAHFALKARTVFGVTVSVITWVAVFASLPGLIIFTSQKEGLHYTC 178
Qy      79 SSHFPYSQYQFWNCFOTLKI-----HORVHGG----- 105
Db      179 SSHFPYSQYQFWNCFOTLKIIVILGVLPFLVMVICYGILTLRCRNKQGRHRAVLIF 238
Qy      106 -----GGSYKCGLC-----QEFPGLNCCSSNRLDGQRV 135
Db      239 TIMIVYFLFWAPYINIVLLNTQEEFPGLNCCSSNRLDQAMQV 281

```

Search completed: October 21, 2003, 17:50:54
Job time : 58.0712 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 17:49:04 ; Search time 19.1797 Seconds
(without alignments)
304.432 Million cell updates/sec

Title: US-10-057-890A-10

Perfect score: 797

Sequence: 1 MDYQVSSPIYDINYTSEPC.....GLNNCSSNRDLGHQRVHAA 138

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA.*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/PCFUS_COMB.pep.*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	379	47.6	352	3	US-09-087-232A-13
2	379	47.6	352	3	US-08-861-105-14
3	379	47.6	352	3	US-08-575-967A-2
4	379	47.6	352	4	US-08-833-752-5
5	379	47.6	352	4	US-09-502-783A-2
6	379	47.6	352	4	US-09-796-202-1
7	374	46.9	352	3	US-09-045-583-52
8	374	46.9	352	3	US-09-534-185-52
9	373	46.8	352	3	US-08-466-343D-2
10	359	45.0	352	4	US-09-517-605-5
11	272	34.1	354	4	US-08-724-984A-2
12	258	32.4	184	4	US-08-833-752-4
13	258	32.4	215	3	US-09-087-232A-17
14	258	32.4	215	4	US-08-833-752-6
15	182.5	22.9	100	3	US-09-087-232A-15
16	159	19.9	87	3	US-09-087-232A-18
17	148.5	18.6	488	2	US-08-933-750C-17
18	148.5	18.6	488	3	US-09-234-613-17
19	128	16.1	54	4	US-08-833-752-11
20	123	15.4	22	3	US-08-861-105-4
21	121	15.2	711	2	US-08-820-170A-10
22	121	15.2	711	3	US-09-055-699-10
23	121	15.2	711	3	US-09-273-565-10
24	121	15.2	711	4	US-09-563-538-10
25	121	15.2	711	4	US-09-661-468-10
26	121	15.2	711	4	US-09-976-165-10
27	118.5	14.9	315	1	US-08-253-155A-34

28	118.5	14.9	532	4	US-09-389-956-80	Sequence 80, Appl
29	117	14.7	457	4	US-09-389-956-68	Sequence 68, Appl
30	117	14.7	647	4	US-09-389-956-6	Sequence 6, Appl
31	114.5	14.4	543	4	US-09-362-123A-4	Sequence 4, Appl
32	114.5	14.4	675	1	US-08-317-522A-9	Sequence 9, Appl
33	114.5	14.4	675	1	US-08-439-818A-9	Sequence 9, Appl
34	114.5	14.4	675	2	US-08-751-965-9	Sequence 9, Appl
35	114.5	14.4	675	2	US-08-738-375-9	Sequence 9, Appl
36	114.5	14.4	675	2	US-08-728-626-9	Sequence 9, Appl
37	114.5	14.4	675	3	US-08-808-599A-9	Sequence 9, Appl
38	114	14.3	21	3	US-08-907-468-11	Sequence 11, Appl
39	113.5	14.2	462	3	US-08-486-099-117	Sequence 117, App
40	113.5	14.2	462	3	US-08-360-107A-127	Sequence 127, App
41	113.5	14.2	462	3	US-08-484-223B-117	Sequence 117, App
42	113.5	14.2	462	3	US-08-919-597-117	Sequence 117, App
43	113.5	14.2	462	3	US-08-475-668A-117	Sequence 117, App
44	113.5	14.2	462	3	US-08-485-551A-117	Sequence 117, App
45	113.5	14.2	462	3	US-08-471-913A-117	Sequence 117, App

ALIGNMENTS

RESULT 1

US-09-087-232A-13
; Sequence 13, Application US/09087232A
; Patent No. 6153431
; GENERAL INFORMATION:
; APPLICANT: Quillent et al.
; TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS CO-RECEPTOR
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Baker & Botts, L.L.P. attn. Lisa Kole
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/087,232A
; FILING DATE: 28 MAY 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/048,057
; FILING DATE: 30 MAY 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: KOLE, LISA B.
; REGISTRATION NUMBER: 35,225
; REFERENCE/DOCKET NUMBER: AP 31115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 408-2628
; TELEFAX: (212) 765-2519
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-087-232A-13

Query Match 47.6%; Score 379; DB 3; Length 352;

Best Local Similarity 35.0%; Pred. No. 1.4e-34;

Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

QY 1 MDYQVSSPIYDINYTSEPCOKINVKQIAA----- 30

DB 1 MDYQVSSPIYDINYTSEPCOKINVKQIAARLLPPLYSLVFIFGVGNMLVILINCKR 60

QY 31 -----YKGLC-----AAQWDFGNTMCQHQRVHGHHSYKCG--- 64
Db 61 LKSWTDIYLLNLALSDFLLTVFPFAHYAAQWDFGNTMC--QLLTGLYFIFGFSGIF 118
QY 65 -----YKGLC-----LCTRSQKGLHYTC 78
Db 119 ILLTIDRYLAVHAFKARTVTFTGVVTSVITWVAVFASLPGLIIFTRSQKGLHYTC 178
QY 79 SSHFPYSQYQFQKQFQTLKI-----HORVHG----- 105
Db 179 SSHFPYSQYQFQKQFQTLKIIVLGLVPLLVVICYSGLKTLRLCRNKKRRAVRLIF 238
QY 106 -----GSKYKGLC-----QEFFGLNCCSSNRLDGHQRV 135
Db 239 TIMIVYFLWAPYNIIVLLNTFQEFFGLNCCSSNRLDQAMQV 281

RESULT 2
US-08-861-105-14
; Sequence 14, Application US/08861105
; Patent No. 6258527
; GENERAL INFORMATION:
; APPLICANT: LITTMAN, DAN R.
; APPLICANT: DENG, HONGKUI
; APPLICANT: ELLMEIER, WILFRIED
; APPLICANT: LANDAU, NATHANIEL R.
; APPLICANT: LIU, RONG
; TITLE OF INVENTION: G-COUPLED RECEPTORS ASSOCIATED WITH
; TITLE OF INVENTION: MACROPHAGE-TROPIC HIV, AND DIAGNOSTIC AND THERAPEUTIC
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave., Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/861.105
; FILING DATE:
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/666,020
; FILING DATE: 19-JUN-1996
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/227,319
; FILING DATE: 13-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1049-1-004 N1
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens

US-08-861-105-14
Query Match 47.6%; Score 379; DB 3; Length 352;
Best Local Similarity 35.0%; Pred. No. 1.4e-34;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;
QY 1 MDQVSSPIYDINVTSEPCQKINVQIAA----- 30
Db 1 MDQVSSPIYDINVTSEPCQKINVQIAAALPPLYSLVFIFGVGNMLVILILNCKR 60
QY 31 -----YKGLC-----AAQWDFGNTMCQHQRVHGHHSYKCG--- 64
Db 61 LKSWTDIYLLNLALSDFLLTVFPFAHYAAQWDFGNTMC--QLLTGLYFIFGFSGIF 118
QY 65 -----YKGLC-----LCTRSQKGLHYTC 78
Db 119 ILLTIDRYLAVHAFKARTVTFTGVVTSVITWVAVFASLPGLIIFTRSQKGLHYTC 178
QY 79 SSHFPYSQYQFQKQFQTLKI-----HORVHG----- 105
Db 179 SSHFPYSQYQFQKQFQTLKIIVLGLVPLLVVICYSGLKTLRLCRNKKRRAVRLIF 238
QY 106 -----GSKYKGLC-----QEFFGLNCCSSNRLDGHQRV 135
Db 239 TIMIVYFLWAPYNIIVLLNTFQEFFGLNCCSSNRLDQAMQV 281
RESULT 3
US-08-575-967A-2
; Sequence 2, Application US/08575967A
; Patent No. 6265184
; GENERAL INFORMATION:
; APPLICANT: Gray et al.
; TITLE OF INVENTION: Chemokine Receptor Materials and Methods
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/575,967A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6265184and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 32918
; TELEPHONE: 206-485-1900
; TELEFAX: 206-485-1662
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: /= "88C amino acid sequence"
US-08-575-967A-2
Query Match 47.6%; Score 379; DB 3; Length 352;
Best Local Similarity 35.0%; Pred. No. 1.4e-34;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

QY 1 MDQVSSPIYDINYITSEPCQKINVKQIAA----- 30
DB 1 MDQVSSPIYDINYITSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
QY 31 -----YKGLC-----AAAQWDFGNTMCHORVHGHHSYKCG--- 64
DB 61 LKSMTDIYLLNLAIISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIFGFSGIF 118
QY 65 -----YKGLC-----LCTRSQEGHLHYTC 78
DB 119 ILLITIDRYLAVHAFALKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGHLHYTC 178
QY 79 SSHFPYSQYQFWKNFQTLKI-----HORVHGG----- 105
DB 179 SSHFPYSQYQFWKNFQTLKIIVILGLVPLLVVICYSGLIKTLRCRNEKRRHRAVRLIF 238
QY 106 -----GGSYKGLC-----QEFFGLNCCSSNRLDGHQV 135
DB 239 TIMIVYFLFWAPYINIVLLNTFQEFFGLNCCSSNRLDQAMQV 281
RESULT 4
US-08-833-752-5
; Sequence 5, Application US/08833752
; Patent No. 6448375
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; APPLICANT: PARMENTIER, MARC
; APPLICANT: VASSART, GILBERT
; APPLICANT: LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knocbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,752
; FILING DATE: 9-APR-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-833-752-5

Query Match 47.6%; Score 379; DB 4; Length 352;
Best Local Similarity 35.0%; Pred. No. 1.4e-34;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;
QY 1 MDQVSSPIYDINYITSEPCQKINVKQIAA----- 30
DB 1 MDQVSSPIYDINYITSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
QY 31 -----YKGLC-----AAAQWDFGNTMCHORVHGHHSYKCG--- 64
DB 61 LKSMTDIYLLNLAIISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIFGFSGIF 118

QY 65 -----LCTRSQEGHLHYTC 78
DB 119 ILLITIDRYLAVHAFALKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGHLHYTC 178
QY 79 SSHFPYSQYQFWKNFQTLKI-----HORVHGG----- 105
DB 179 SSHFPYSQYQFWKNFQTLKIIVILGLVPLLVVICYSGLIKTLRCRNEKRRHRAVRLIF 238
QY 106 -----GGSYKGLC-----QEFFGLNCCSSNRLDGHQV 135
DB 239 TIMIVYFLFWAPYINIVLLNTFQEFFGLNCCSSNRLDQAMQV 281
RESULT 5
US-09-502-783A-2
; Sequence 2, Application US/09502783A
; Patent No. 6511826
; GENERAL INFORMATION:
; APPLICANT: Li, Yi
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein Chemokine Receptor (C
; FILE REFERENCE: 1488.1150006
; CURRENT APPLICATION NUMBER: US/09/502,783A
; PRIOR FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: 08/466,343
; PRIOR FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-502-783A-2
Query Match 47.6%; Score 379; DB 4; Length 352;
Best Local Similarity 35.0%; Pred. No. 1.4e-34;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;
QY 1 MDQVSSPIYDINYITSEPCQKINVKQIAA----- 30
DB 1 MDQVSSPIYDINYITSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
QY 31 -----YKGLC-----AAAQWDFGNTMCHORVHGHHSYKCG--- 64
DB 61 LKSMTDIYLLNLAIISDLFFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIFGFSGIF 118
QY 65 -----LCTRSQEGHLHYTC 78
DB 119 ILLITIDRYLAVHAFALKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGHLHYTC 178
QY 79 SSHFPYSQYQFWKNFQTLKI-----HORVHGG----- 105
DB 179 SSHFPYSQYQFWKNFQTLKIIVILGLVPLLVVICYSGLIKTLRCRNEKRRHRAVRLIF 238
QY 106 -----GGSYKGLC-----QEFFGLNCCSSNRLDGHQV 135
DB 239 TIMIVYFLFWAPYINIVLLNTFQEFFGLNCCSSNRLDQAMQV 281

RESULT 6
US-09-796-202-1
; Sequence 1, Application US/09796202
; Patent No. 6548636
; GENERAL INFORMATION:
; APPLICANT: Olson, William
; APPLICANT: Dragic, Tatjana
; TITLE OF INVENTION: SULFATED CCR5 PEPTIDES FOR HIV-1 INFECTION
; FILE REFERENCE: 2048/61010/JPW/SHS
; CURRENT APPLICATION NUMBER: US/09/796,202
; CURRENT FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.0

```
; SEQ ID NO 1
; LENGTH: 352
; TYPE: PRT
; ORGANISM: human
US-09-796-202-1

Query Match      47.6%; Score 379; DB 4; Length 352;
Best Local Similarity 35.0%; Pred. No. 1.4e-34;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

QY 1 MDQVSSPIYDINVTSEPCQKINVKQIAA----- 30
DB 1 MDQVSSPIYDINVTSEPCQKINVKQIAARLLPPLYSLVFIFGFGVGNMLVILINCKR 60
QY 31 -----YKGLC----- 64
DB 61 LKSMTDIYLLNLAIISDLFFLLTPVFWAHYAAQWDFGNTMC--QLLTGLYFIQFFSGIFF 118
QY 65 -----LCTRSQKEGLHYTC 78
DB 119 IILITDRIYLAIVHVPALKARTVTFGVTSVITWVAVPASLPGIIFTRSQKEGLHYTC 178
QY 79 SSHFPYSQYQFKNFQTLKI----- 105
DB 179 SSHFPYSQYQFKNFQTLKIVILGLVPLLVVICYSGLILKTLRCRNEKGRHRAVLIF 238
QY 106 -----GSKYKGLC----- 135
DB 239 TIMIVYFLFWAPYINIVLLNTFOEFFGLNCCSSNRLDQAMQV 281

RESULT 7
US-09-045-583-52
; Sequence 52, Application US/09045583
; Patent No. 6287805
; GENERAL INFORMATION:
; APPLICANT: Graham, Gerard J. et al.
; TITLE OF INVENTION: No. 6287805el Molecules of the G Protein-Coupled
; RECEPTOR SUPERFAMILY AND USES
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/045,583
; FILING DATE: 20-MAR-98
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragoras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: MNI-044
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)742-4214
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-09-045-583-52

Query Match      46.9%; Score 374; DB 3; Length 352;
Best Local Similarity 34.6%; Pred. No. 5.1e-34;
Matches 98; Conservative 8; Mismatches 27; Indels 150; Gaps 7;

QY 1 MDQVSSPIYDINVTSEPCQKINVKQIAA----- 30
DB 1 MDQVSSPIYDINVTSEPCQKINVKQIAARLLPPLYSLVFIFGFGVGNMLVILINCKR 60
QY 31 -----YKGLC----- 64
DB 61 LKSMTDIYLLNLAIISDLFFLLTPVFWAHYAAQWDFGNTMC--QLLTGLYFIQFFSGIFF 118
QY 65 -----LCTRSQKEGLHYTC 78
DB 119 IILITDRIYLAIVHVPALKARTVTFGVTSVITWVAVPASLPGIIFTRSQKEGLHYTC 178
QY 79 SSHFPYSQYQFKNFQTLKI----- 105
DB 179 SSHFPYSQYQFKNFQTLKIVILGLVPLLVVICYSGLILKTLRCRNEKGRHRAVLIF 238
QY 106 -----GSKYKGLC----- 135
DB 239 TIMIVYFLFWAPYINIVLLNTFOEFFGLNCCSSNRLDQAMQV 281

RESULT 8
US-09-534-185-52
; Sequence 52, Application US/09534185
; Patent No. 6403767
; GENERAL INFORMATION:
; APPLICANT: Graham, Gerard J. et al.
; TITLE OF INVENTION: No. 6403767el Molecules of the G Protein-Coupled
; RECEPTOR SUPERFAMILY AND USES
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/534,185
; FILING DATE: 24-Mar-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/045,583
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragoras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: MNI-044
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 52:
US-09-534-185-52

Query Match      46.9%; Score 374; DB 4; Length 352;
```


Best Local Similarity 34.6%; Pred. No. 5.1e-34;
Matches 98; Conservative 8; Mismatches 27; Indels 150; Gaps 7;

Qy 1 MDYQVSSPIVDIYNTSEPCQKINVKQIAA----- 30
Db 1 MDYQVSSPIVDIYNTSEPCQKINVKQIAARLLPPLYSLVIFGPGVGNMLVILINCKR 60
Qy 31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
Db 61 LKSMTDIYLLNLALSDFLLTVFPFMAHYAAQWDFGNTMC--QLLTGLYFIFGFSGIFF 118
Qy 65 -----LCTRSQKEGLHYTC 78
Db 119 ILLTIDRYLAIVHAVFALKARTVTFGWVTSVITWVAVFASLPGLIIFTRSQEGLHYTC 178
Qy 79 SSHFPYSQYQFWKQFOTLKI-----HQRVHGG----- 105
Db 179 SSHFPYSQYQFWKQFOTLKI-----HQRVHGG----- 105
Qy 106 -----GGSYKCGLC-----QEFFGLNCCSSNRLDGHQV 135
Db 239 TIMIVYFLWAPYINVLNLTFFQEFFGLNCCSSNRLDQAMQV 281

RESULT 9
US-08-466-343D-2
; Sequence 2, Application US/08466343D
; Patent No. 6025154
; GENERAL INFORMATION:
; APPLICANT: LI, Yi
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING HUMAN G-PROTEIN
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,343D
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1150000/EKS/KLM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2500
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-466-343D-2

Query Match 46.8%; Score 373; DB 3; Length 352;
Best Local Similarity 34.6%; Pred. No. 6.6e-34;
Matches 98; Conservative 7; Mismatches 28; Indels 150; Gaps 7;

Qy 1 MDYQVSSPIVDIYNTSEPCQKINVKQIAA----- 30
Db 1 MDYQVSSPIVDIYNTSEPCQKINVKQIAARLLPPLYSLVIFGPGVGNMLVILINCKR 60
Qy 31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64

Query Match 45.0%; Score 359; DB 4; Length 352;
Best Local Similarity 33.9%; Pred. No. 2.4e-32;
Matches 96; Conservative 8; Mismatches 29; Indels 150; Gaps 7;

Qy 1 MDYQVSSPIVDIYNTSEPCQKINVKQIAA----- 30
Db 1 MDYQVSSPTDIDYDTSQPCQKINVKQIAARLLPPLYSLVIFGPGVGNMLVILINCKR 60
Qy 31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
Db 61 LKSMTDIYLLNLALSDFLLTVFPFMAHYAAQWDFGNTMC--QLLTGLYFIFGFSGIFF 118
Qy 65 -----LCTRSQKEGLHYTC 78
Db 119 ILLTIDRYLAIVHAVFALKARTVTFGWVTSVITWVAVFASLPGLIIFTRSQEGLHYTC 178
Qy 79 SSHFPYSQYQFWKQFOTLKI-----HQRVHGG----- 105
Db 179 SSHFPYSQYQFWKQFOTLKI-----HQRVHGG----- 105
Qy 106 -----GGSYKCGLC-----QEFFGLNCCSSNRLDGHQV 135
Db 239 TIMIVYFLWAPYINVLNLTFFQEFFGLNCCSSNRLDQAMQV 281

RESULT 11
US-08-724-984A-2
; Sequence 2, Application US/08724984A
; Patent No. 6388055
; GENERAL INFORMATION:
; APPLICANT: Derk Bergsma, Mary Brawner, and Usman Shabon
; TITLE OF INVENTION: No. 6388055el Mouse Genomic Clone of the CC-
; TITLE OF INVENTION: CKR5 Receptor
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:

Db 61 LESMTDIYLLNLALSDFLLTVFPFMAHYAAQWDFGNTMC--QLLTGLYFIFGFSGIFF 118
Qy 65 -----LCTRSQKEGLHYTC 78
Db 119 ILLTIDRYLAIVHAVFALKARTVTFGWVTSVITWVAVFASLPGLIIFTRSQEGLHYTC 178
Qy 79 SSHFPYSQYQFWKQFOTLKI-----HQRVHGG----- 105
Db 179 SSHFPYSQYQFWKQFOTLKI-----HQRVHGG----- 105
Qy 106 -----GGSYKCGLC-----QEFFGLNCCSSNRLDGHQV 135
Db 239 TIMIVYFLWAPYINVLNLTFFQEFFGLNCCSSNRLDQAMQV 281

RESULT 10
US-09-517-605-5
; Sequence 5, Application US/09517605
; Patent No. 6391567
; GENERAL INFORMATION:
; APPLICANT: Littman, Dan R.
; APPLICANT: Kwon, Douglas S.
; APPLICANT: van Kooyk, Yvette
; APPLICANT: Geijtenbeck, Theo
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 1049-1-017
; CURRENT APPLICATION NUMBER: US/09/517,605
; CURRENT FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-517-605-5

Query Match 45.0%; Score 359; DB 4; Length 352;
Best Local Similarity 33.9%; Pred. No. 2.4e-32;
Matches 96; Conservative 8; Mismatches 29; Indels 150; Gaps 7;

Qy 1 MDYQVSSPIVDIYNTSEPCQKINVKQIAA----- 30
Db 1 MDYQVSSPTDIDYDTSQPCQKINVKQIAARLLPPLYSLVIFGPGVGNMLVILINCKR 60
Qy 31 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
Db 61 LKSMTDIYLLNLALSDFLLTVFPFMAHYAAQWDFGNTMC--QLLTGLYFIFGFSGIFF 118
Qy 65 -----LCTRSQKEGLHYTC 78
Db 119 ILLTIDRYLAIVHAVFALKARTVTFGWVTSVITWVAVFASLPGLIIFTRSQEGLHYTC 178
Qy 79 SSHFPYSQYQFWKQFOTLKI-----HQRVHGG----- 105
Db 179 SSHFPYSQYQFWKQFOTLKI-----HQRVHGG----- 105
Qy 106 -----GGSYKCGLC-----QEFFGLNCCSSNRLDGHQV 135
Db 239 TIMIVYFLWAPYINVLNLTFFQEFFGLNCCSSNRLDQAMQV 281

RESULT 11
US-08-724-984A-2
; Sequence 2, Application US/08724984A
; Patent No. 6388055
; GENERAL INFORMATION:
; APPLICANT: Derk Bergsma, Mary Brawner, and Usman Shabon
; TITLE OF INVENTION: No. 6388055el Mouse Genomic Clone of the CC-
; TITLE OF INVENTION: CKR5 Receptor
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:

ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road, P.O. Box 1539
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
COMPUTER: IBM 486
OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
SOFTWARE: MICROSOFT WORD
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,984A
FILING DATE: October 3, 1996
CLASSIFICATION: 800
PRIOR APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
NAME: William T. Han
REGISTRATION NUMBER: 34,344
REFERENCE/DOCKET NUMBER: ATGS0023
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610 270 5024
TELEFAX: 610 270 5090
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 354
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-724-984A-2

Query Match 34.1%; Score 272; DB 4; Length 354;
Best Local Similarity 27.9%; Pred. No. 1.3e-22;
Matches 80; Conservative 12; Mismatches 29; Indels 166; Gaps 8;
QY 1 MDYQVSSP--YDINYYTSEPCQKINVKQIAA----- 30
DB 1 MDQGSVPYTYDIDYMGAPCQKINVKQIAAQLLPYLSLVFIFGAGNMVFLILISC 60
QY 31 -----YKGLC----- 30
DB 61 KKLKSVTDIYLLNLAIISDLFLTLFPWYAAAEWIFGNIMC--KVFTGVHIGYFGGI 118
QY 65 -----LCTRSQKEGLHY 76
DB 119 FFIILLTDIYLAIVHAVFALKVTTVNFVITSVTVVAVFASLPFELIFTSQKEGHY 178
QY 77 TCSHPFPYQYQWKNQFOLK-----IHQVHGGGYSYK----- 111
DB 179 TCSHPFPYQYHFKWSFOTLKMVILSLILPLVMIICYSGLIHT-----LFCRCNEKK 231
QY 112 -----GLCOEFFGLNCCSSNRILD 130
DB 232 RHRVRLIFAIMIVYFLFWTPYINVILLTTFEFGFLNCCSSNRILD 278

RESULT 12
US-08-833-752-4
Sequence 4, Application US/08833752
Patent No. 6448375
GENERAL INFORMATION:
APPLICANT: SAMSON, MICHEL
APPLICANT: PARMENTIER, MARC
APPLICANT: VASSART, GILBERT
APPLICANT: LIBERT, FREDERICK
TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
TITLE OF INVENTION: AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach

STATE: CA
COUNTRY: U.S.A.
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,752
FILING DATE: 9-APR-1997
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 184 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-833-752-4
Query Match 32.4%; Score 258; DB 4; Length 184;
Best Local Similarity 34.9%; Pred. No. 2.3e-21;
Matches 65; Conservative 4; Mismatches 13; Indels 104; Gaps 4;
QY 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
DB 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAAQLLPYLSLVFIFGVGNMVLILINCKR 60
QY 31 -----YKGLC----- 30
DB 61 LKSMTDIYLLNLAIISDLFLTLFPWYAAAEWIFGNIMC--OLLTGLYFTIGFSGIIF 118
QY 65 -----LCTRSQKEGLHYTC 78
DB 119 IILLTDIYLAIVHAVFALKARTVTFGVITSVITVAVFASLPGLIFTRSQKEGLHYTC 178
QY 79 SSHPPY 84
DB 179 SSHPPY 184

RESULT 13
US-09-087-232A-17
Sequence 17, Application US/09087232A
Patent No. 6153431
GENERAL INFORMATION:
APPLICANT: Quillient et al.
TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS CO-RECEPTOR
TITLE OF INVENTION: VARIANTS ASSOCIATED WITH RESISTANCE TO VIRUS INFECTION.
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Baker & Botts, L.L.P. attn. Lisa Kole
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/087,232A
FILING DATE: 28 MAY 1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/048,057
FILING DATE: 30 MAY 1997

```
; ATTORNEY/AGENT INFORMATION:
; NAME: KOLE, LISA B.
; REGISTRATION NUMBER: 35,225
; REFERENCE/DOCKET NUMBER: AP 31115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 408-2628
; TELEFAX: (212) 765-2519
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-087-232A-17

Query Match 32.4%; Score 258; DB 3; Length 215;
Best Local Similarity 34.9%; Pred. No. 2.8e-21;
Matches 65; Conservative 4; Mismatches 13; Indels 104; Gaps 4;

QY 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
DB 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLSLVIFGFVGNMLVILILINCKR 60
QY 31 -----YKGLC-----AAAQWDFGNTMCOHQRVHGHHSYKCG--- 64
DB 61 LKSMTDIYLLNLALSDFLLTVPFWAHYAAAQWDFGNTMC--QLLTGLYFIFGFFSGIFF 118
QY 65 -----LCTRSQKEGLHYTC 78
DB 119 ILLTIDRYLAVHVAFFALKARTVTFGVVTSVITWVAVFASPLPGIIFTRSQKEGLHYTC 178
QY 79 SSHFPY 84
DB 179 SSHFPY 184

RESULT 14
US-08-833-752-6
; Sequence 6, Application US/08833752
; Patent No. 6448375
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; APPLICANT: PARMENTIER, MARC
; APPLICANT: VASSART, GILBERT
; APPLICANT: LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; TITLE OF INVENTION: AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,752
; FILING DATE: 9-APR-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; ATTORNEY/AGENT INFORMATION:
; NAME: KOLE, LISA B.
; REGISTRATION NUMBER: 35,225
; REFERENCE/DOCKET NUMBER: AP 31115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 408-2628
; TELEFAX: (212) 765-2519
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 100 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-087-232A-15

Query Match 22.9%; Score 182.5; DB 3; Length 100;
Best Local Similarity 43.0%; Pred. No. 3.2e-13;
Matches 43; Conservative 0; Mismatches 4; Indels 53; Gaps 2;

QY 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
```

Db 1 MDYQVSSPIYDINYTTSEPCQKINVKQIAARLLPLYSLVFIFGFVGNMLVILINCKR 60
QY 31 -----YKGLC-----AAQWDFGNTM 47
Db 61 LKSMTDIYLLNLAIISDLFLLTVPFWAHYAAQWDFGNTM 100

Search completed: October 21, 2003, 17:55:23
Job time : 21.1797 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 17:53:49 ; Search time 42.1017 Seconds
(without alignments)
548.899 Million cell updates/sec

Title: US-10-057-890A-10

Perfect score: 797

Sequence: 1 MDYQVSSPIVDINYITSEPC.....GLNCCSSNRLDGHQVHAA 138

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 629382 seqs, 167460630 residues

Total number of hits satisfying chosen parameters: 629382

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	797	100.0	138	15	US-10-057-890A-10
2	797	100.0	157	15	Sequence 31, Appl
3	379	47.6	352	9	US-09-725-285-2
4	379	47.6	352	9	US-09-759-841-2
5	379	47.6	352	9	US-09-779-879A-22
6	379	47.6	352	9	US-09-779-880A-22
7	379	47.6	352	9	US-09-813-653-15
8	379	47.6	352	9	US-09-813-653-17
9	379	47.6	352	9	US-09-796-202-1
10	379	47.6	352	9	US-09-195-662A-2
11	379	47.6	352	9	US-09-339-912A-2
12	379	47.6	352	10	US-09-938-719-5
13	379	47.6	352	10	US-09-938-226-5
14	379	47.6	352	10	US-09-938-703-5
15	379	47.6	352	10	US-09-502-783A-2

16	379	47.6	352	11	US-09-734-221A-14	Sequence 14, Appl
17	379	47.6	352	12	US-10-323-314-1	Sequence 1, Appl
18	379	47.6	352	12	US-10-072-301-1	Sequence 1, Appl
19	379	47.6	352	12	US-10-071-866-1	Sequence 1, Appl
20	379	47.6	352	12	US-10-135-839-22	Sequence 22, Appl
21	379	47.6	352	12	US-10-239-423-67	Sequence 67, Appl
22	379	47.6	352	12	US-10-439-845-4	Sequence 4, Appl
23	379	47.6	352	14	US-10-106-623-2	Sequence 2, Appl
24	379	47.6	352	15	US-10-232-682-2	Sequence 2, Appl
25	379	47.6	352	15	US-10-086-814-1	Sequence 1, Appl
26	379	47.6	352	15	US-10-067-800-22	Sequence 22, Appl
27	379	47.6	352	15	US-10-290-058A-6	Sequence 6, Appl
28	379	47.6	352	15	US-10-225-567A-352	Sequence 352, Appl
29	374	46.9	352	12	US-10-164-649-52	Sequence 52, Appl
30	374	46.9	352	12	US-10-439-845-2	Sequence 2, Appl
31	373	46.8	352	9	US-09-779-879A-2	Sequence 2, Appl
32	373	46.8	352	9	US-09-779-880A-2	Sequence 2, Appl
33	373	46.8	352	12	US-10-135-839-2	Sequence 2, Appl
34	373	46.8	352	15	US-10-067-800-2	Sequence 2, Appl
35	363	45.5	352	14	US-10-106-623-20	Sequence 20, Appl
36	258	32.4	184	10	US-09-938-719-4	Sequence 4, Appl
37	258	32.4	184	10	US-09-939-226-4	Sequence 4, Appl
38	258	32.4	184	10	US-09-938-703-4	Sequence 4, Appl
39	258	32.4	215	10	US-09-938-719-6	Sequence 6, Appl
40	258	32.4	215	10	US-09-939-226-6	Sequence 6, Appl
41	258	32.4	215	10	US-09-938-703-6	Sequence 6, Appl
42	183.5	23.0	332	12	US-10-095-876A-2	Sequence 2, Appl
43	182	22.8	32	12	US-10-072-301-9	Sequence 9, Appl
44	182	22.8	32	12	US-10-071-866-9	Sequence 9, Appl
45	182	22.8	32	12	US-10-239-423-42	Sequence 42, Appl

ALIGNMENTS

RESULT 1
US-10-057-890A-10
; Sequence 10, Application US/10057890A
; Publication No. US20030044901A1
; GENERAL INFORMATION:
; APPLICANT: Coleman, Timothy
; APPLICANT: Mansfield, Brian
; TITLE OF INVENTION: Scaffold
; FILE OF INVENTION: of Using the Same.
; FILE REFERENCE: PF537
; CURRENT APPLICATION NUMBER: US/10/057,890A
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: 60/265,782
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,858
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 10
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-890A-10

Query Match	100.0%	Score 797;	DB 15;	Length 138;
Best Local Similarity	100.0%	Pred. No. 2.7e-71;		
Matches 138;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MDYQVSSPIVDINYITSEPCQKINVKQIAAYKGLCAAQWDFGNTMCOHQRVHGHHS	60	
Db	1	MDYQVSSPIVDINYITSEPCQKINVKQIAAYKGLCAAQWDFGNTMCOHQRVHGHHS	60	
Qy	61	YKCGLCSTRSQKGLHYTCSSHPFYQFWNFQTLKHQRVHGHGSYKCGLCQBFPG	120	
Db	61	YKCGLCSTRSQKGLHYTCSSHPFYQFWNFQTLKHQRVHGHGSYKCGLCQBFPG	120	
Qy	121	NNCCSSNRLDGHQVHAA	138	
Db	121	NNCCSSNRLDGHQVHAA	138	

```
RESULT 2
US-10-057-890A-31
; Sequence 31, Application US/10057890A
; Publication No. US20030044901A1
; GENERAL INFORMATION:
; APPLICANT: Coleman, Timothy
; APPLICANT: Mansfield, Brian
; TITLE OF INVENTION: Scaffold Fusion Polypeptides, Composition for Making the Same, and
; TITLE OF INVENTION: of Using the Same.
; FILE REFERENCE: PF537
; CURRENT APPLICATION NUMBER: US/10/057,890A
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: 60/265,782
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,858
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 31
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-890A-31

Query Match 100.0%; Score 797; DB 15; Length 157;
Best Local Similarity 100.0%; Pred. No. 3e-71;
Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDYQVSSPIVDINYYTSEPCQKINVKQIAAAYKGLGCLCAAQWDFGNTMCOHORVHGHHHS 60
Db 20 MDYQVSSPIVDINYYTSEPCQKINVKQIAAAYKGLGCLCAAQWDFGNTMCOHORVHGHHHS 79
Qy 61 YKGLCTRQKEGLHYTCSSHPFSQYQFKNFQTLKIHORVHGGGSGYKGLCQEFFGL 120
Db 80 YKGLCTRQKEGLHYTCSSHPFSQYQFKNFQTLKIHORVHGGGSGYKGLCQEFFGL 139
Qy 121 NNCSSNRLDGHQVHAA 138
Db 140 NNCSSNRLDGHQVHAA 157

RESULT 3
US-09-725-285-2
; Sequence 2, Application US/09725285
; Patent No. US20010000241A1
; GENERAL INFORMATION:
; APPLICANT: Li, Yi
; APPLICANT: Ruben, Steven, M.
; TITLE OF INVENTION: Antibodies to Human G-Protein Chemokine Receptor HDGNR10
; TITLE OF INVENTION: (CCR5 Receptor)
; FILE REFERENCE: 1488.1150003
; CURRENT APPLICATION NUMBER: US/09/725,285
; CURRENT FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 09/339,912
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 09/195,662
; PRIOR FILING DATE: 1998-11-18
; PRIOR APPLICATION NUMBER: 08/466,343
; PRIOR FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Artificial Sequence: Genomic
; FEATURE:
; OTHER INFORMATION: Deduced Amino Acid Sequence
US-09-725-285-2

Query Match 47.6%; Score 379; DB 9; Length 352;
Best Local Similarity 35.0%; Pred. No. 1.4e-29;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;
```

```
Qy 1 MDYQVSSPIVDINYYTSEPCQKINVKQIAA----- 30
Db 1 MDYQVSSPIVDINYYTSEPCQKINVKQIAARLLPPLYSLVFIQFVGNMLVILILINCKR 60
Qy 31 -----YKGLGCL-----AAAQWDFGNTMCOHORVHGHHHSYKCG--- 64
Db 61 LKSMTDIYLLNLAIISDLFLLTVPFPAHYAAQWDFGNTMC--QLLTGLYFIQFSGIFF 118
Qy 65 -----LCTRQKEGLHYTC 78
Db 119 IILLTIDRYLAVVHAFALKARTVTGVTTSVITWVAVFASLPGIIFTRQKEGLHYTC 178
Qy 79 SSHFPYSQYQFKNFQTLKI-----HORVHG----- 105
Db 179 SSHFPYSQYQFKNFQTLKIIVILGLVPLLVVICYSGLKTLRCRNEKKRRAVRLIF 238
Qy 106 -----GGSYKGLC-----QEFFGLNCSNRLDGHQV 135
Db 239 TIMIVYFLWAPYNIIVLLNTFOEFFGLNCSNRLDQAMQV 281

RESULT 4
US-09-759-841-2
; Sequence 2, Application US/09759841
; Patent No. US20010039026A1
; GENERAL INFORMATION:
; APPLICANT: Rickett, Graham A
; APPLICANT: Dobbs, Susan
; APPLICANT: Petros, Manousos
; TITLE OF INVENTION: Assay Method
; FILE REFERENCE: PC10348APME
; CURRENT APPLICATION NUMBER: US/09/759,841
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: GB 0006661.9
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: GB 0006663.5
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: GB 0006659.3
; PRIOR FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-759-841-2

Query Match 47.6%; Score 379; DB 9; Length 352;
Best Local Similarity 35.0%; Pred. No. 1.4e-29;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

Qy 1 MDYQVSSPIVDINYYTSEPCQKINVKQIAA----- 30
Db 1 MDYQVSSPIVDINYYTSEPCQKINVKQIAARLLPPLYSLVFIQFVGNMLVILILINCKR 60
Qy 31 -----YKGLGCL-----AAAQWDFGNTMCOHORVHGHHHSYKCG--- 64
Db 61 LKSMTDIYLLNLAIISDLFLLTVPFPAHYAAQWDFGNTMC--QLLTGLYFIQFSGIFF 118
Qy 65 -----LCTRQKEGLHYTC 78
Db 119 IILLTIDRYLAVVHAFALKARTVTGVTTSVITWVAVFASLPGIIFTRQKEGLHYTC 178
Qy 79 SSHFPYSQYQFKNFQTLKI-----HORVHG----- 105
Db 179 SSHFPYSQYQFKNFQTLKIIVILGLVPLLVVICYSGLKTLRCRNEKKRRAVRLIF 238
Qy 106 -----GGSYKGLC-----QEFFGLNCSNRLDGHQV 135
Db 239 TIMIVYFLWAPYNIIVLLNTFOEFFGLNCSNRLDQAMQV 281
```



```
QY 79 SSHFPYSQYQFKNFQTLKI-----HORVHGG----- 105
DQ 179 SSHFPYSQYQFKNFQTLKIIVLGLVPLLVWVICYSGILKTLRLCRNEKRRHRAVRLIF 238
QY 106 -----GSKYKGLC-----QEFFGLNCCSSNRLDGHQV 135
DQ 239 TIMIVYFLFWAPYINIVLLNTFQEFFGLNCCSSNRLDQAMQV 281

RESULT 8
US-09-813-653-17
; Sequence 17, Application US/09813653
; Patent No. US20020064770A1
; GENERAL INFORMATION:
; APPLICANT: Nestor, John
; APPLICANT: Wilson, Carol
; APPLICANT: See, Raymond
; APPLICANT: Tan Hehir, Christina
; TITLE OF INVENTION: Binding Compounds and Methods For Identifying Binding Compounds
; FILE REFERENCE: CNS-005
; CURRENT APPLICATION NUMBER: US/09/813,653
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US 60/190,946
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/190,996
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/191,299
; PRIOR FILING DATE: 2000-03-21
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-813-653-17

Query Match 47.6%; Score 379; DB 9; Length 352;
Best Local Similarity 35.0%; Pred. No. 1.4e-29;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

QY 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
DQ 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVIFGFVGNMLVILINCKR 60
QY 31 -----YKCGLC-----AAAQWDFGNTMCOHORVGHGHHHSYKCG--- 64
DQ 61 LKSMTDIYLLNLALISDLFFLLTPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118
QY 65 -----LCTRSQKEGLHYTC 78
DQ 119 ILLTIDRYLAVHVAFAKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178
QY 79 SSHFPYSQYQFKNFQTLKI-----HORVHGG----- 105
DQ 179 SSHFPYSQYQFKNFQTLKIIVLGLVPLLVWVICYSGILKTLRLCRNEKRRHRAVRLIF 238
QY 106 -----GSKYKGLC-----QEFFGLNCCSSNRLDGHQV 135
DQ 239 TIMIVYFLFWAPYINIVLLNTFQEFFGLNCCSSNRLDQAMQV 281

RESULT 9
US-09-796-202-1
; Sequence 1, Application US/09796202
; Patent No. US20020068813A1
; GENERAL INFORMATION:
; APPLICANT: Dragic, Tatjana
; APPLICANT: Olson, William
; TITLE OF INVENTION: SULFATED CCR5 PEPTIDES FOR HIV-1 INFECTION
; FILE REFERENCE: 2048/61010/JPW/SHS
; CURRENT APPLICATION NUMBER: US/09/796,202
; CURRENT FILING DATE: 2001-02-28
```

```
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 352
; TYPE: PRT
; ORGANISM: human
US-09-796-202-1

Query Match 47.6%; Score 379; DB 9; Length 352;
Best Local Similarity 35.0%; Pred. No. 1.4e-29;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

QY 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
DQ 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVIFGFVGNMLVILINCKR 60
QY 31 -----YKCGLC-----AAAQWDFGNTMCOHORVGHGHHHSYKCG--- 64
DQ 61 LKSMTDIYLLNLALISDLFFLLTPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118
QY 65 -----LCTRSQKEGLHYTC 78
DQ 119 ILLTIDRYLAVHVAFAKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178
QY 79 SSHFPYSQYQFKNFQTLKI-----HORVHGG----- 105
DQ 179 SSHFPYSQYQFKNFQTLKIIVLGLVPLLVWVICYSGILKTLRLCRNEKRRHRAVRLIF 238
QY 106 -----GSKYKGLC-----QEFFGLNCCSSNRLDGHQV 135
DQ 239 TIMIVYFLFWAPYINIVLLNTFQEFFGLNCCSSNRLDQAMQV 281

RESULT 10
US-09-195-662A-2
; Sequence 2, Application US/09195662A
; Patent No. US20020076745A1
; GENERAL INFORMATION:
; APPLICANT: Li, Vi
; APPLICANT: Ruben, Steven, M.
; TITLE OF INVENTION: Human G-Protein Chemokine Receptor HDGNR10 (CCR5 Receptor)
; CURRENT APPLICATION NUMBER: US/09/195,662A
; CURRENT FILING DATE: 1998-11-18
; PRIOR APPLICATION NUMBER: 08/466,343
; PRIOR FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Artificial Sequence: Genomic
; FEATURE:
; OTHER INFORMATION: Deduced Amino Acid Sequence
US-09-195-662A-2

Query Match 47.6%; Score 379; DB 9; Length 352;
Best Local Similarity 35.0%; Pred. No. 1.4e-29;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

QY 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
DQ 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVIFGFVGNMLVILINCKR 60
QY 31 -----YKCGLC-----AAAQWDFGNTMCOHORVGHGHHHSYKCG--- 64
DQ 61 LKSMTDIYLLNLALISDLFFLLTPFWAHYAAAQWDFGNTMC--QLLTGLYFIGFFSGIFF 118
QY 65 -----LCTRSQKEGLHYTC 78
DQ 119 ILLTIDRYLAVHVAFAKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178
QY 79 SSHFPYSQYQFKNFQTLKI-----HORVHGG----- 105
```



```

: STREET: 620 Newport Center Drive 16th Floor
: CITY: Newport Beach
: STATE: CA
: COUNTRY: U.S.A.
: ZIP: 92660
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/938,719
: FILING DATE: 24-Aug-2001
: CLASSIFICATION: <Unknown>
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 09/626,939
: FILING DATE: 27-JULY-2000
: ATTORNEY/AGENT INFORMATION:
: NAME: Altman, Daniel E
: REGISTRATION NUMBER: 34,115
: REFERENCE/DOCKET NUMBER: <Unknown>
:
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 352 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
:
: MOLECULE TYPE: protein
:
: SEQUENCE DESCRIPTION: SEQ ID NO: 5:
:
: US-09-938-719-5
:
: Query Match 47.6%; Score 379; DB 10; Length 352;
: Best Local Similarity 35.0%; Pred. No. 1.4e-29;
: Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 0
:
: QY 1 MDYQSSPIYDINYTSEPCKINVKQIAA----- 30
: DB 1 MDYQSSPIYDINYTSEPCKINVKQIAARLLPPLSLVIFFGVGNMLVILINCKR 60
:
: QY 31 -----YKGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 64
: DB 61 LKSMTDIYLLNLAIISDLFLLTVPFMAHYAAAQWDFGNTWC--QLLTGLYFICFFSGIFF 178
:
: QY 65 -----LCTRSQKEGLHYTC 78
: DB 119 IILLTDIYLAVHAFKALKARTVTFGWTSVITWVAVFASLPFGIFTRSQKEGLHYTC 178
:
: QY 79 SSHFPYSQYQFWKQFQTLKI-----HORVHGG----- 105
: DB 179 SSHFPYSQYQFWKQFQTLKIIVLGLVPLVMVICYSGLIKTLRCNEKKRHAVLFI 238
:
: QY 106 -----GSGYKGLC-----QEFGGLNCCSSNRLLDGHQV 135
: DB 239 TIMIVYFLFWAPYINIVLLNTFQEFFGLNCCSSNRLLDAQMV 281
:
: RESULT 13
: US-09-939-226-5
: ; Sequence 5, Application US/09939226
: ; Patent No. US20020110805A1
: ; GENERAL INFORMATION:
: ; APPLICANT: SAMSON, MICHEL
: ; PARMENTIER, MARC
: ; VASSART, GILBERT
: ; LIBERT, FREDERICK
:
: TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
:
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Knobbe, Martens, Olson & Bear
: STREET: 620 Newport Center Drive 16th Floor
: CITY: Newport Beach
: STATE: CA
: COUNTRY: U.S.A.

```

ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/939,226
FILING DATE: 24-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/939,226
FILING DATE: 24-Aug-2001
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER: <Unknown>
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-939-226-5

Query Match 47.6%; Score 379; DB 10; Length 352;
Best Local Similarity 35.0%; Pred. No. 1.4e-29;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

QY 1 MDYQVSSPIVDINYYTSEPCQKINVKQIAA----- 30
DB 1 MDYQVSSPIVDINYYTSEPCQKINVKQIAALLPPLYSLVFIFFGVGNMLVILINCKR 60
QY 31 -----YKGLC-----AAQWDFGNTMCOHQRVHGHHSYKCG--- 64
DB 61 LKSMTDIYLLNLAISDLFLLTVFPFAHYAAAQWDFGNTMC--QLLTGLYFIFFSGIFF 118
QY 65 -----YKGLC-----LCTRSQKEGLHYTC 78
DB 119 ILLTIDRYLAVHVAFAKARTVTFGVTSVITWVAVFASLPGLIFTRSQKEGLHYTC 178
QY 79 SSHFPYSQYQFKNFOTLKI-----HORVHGG----- 105
DB 179 SSHFPYSQYQFKNFOTLKIIVILGLVPLLMVICYSGLKTLRLCRNEKKHRAVRLIF 238
QY 106 -----GGSYKGLC-----QEFFGLNCCSSNRLDGHQRV 135
DB 239 TIMIVYFLFAPYNIIVLLNTFOEFFGLNCCSSNRLDQAMQV 281

RESULT 14
US-09-938-703-5
Sequence 5, Application US/09938703
Patent No. US20020110870A1
GENERAL INFORMATION:
APPLICANT: SAMSON, MICHEL
PARMENTIER, MARC
VASSART, GILBERT
LIBERT, FREDERICK
TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: U.S.A.
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/938,703
FILING DATE: 24-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/626,939
FILING DATE: 2000-07-27
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER: <Unknown>
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-938-703-5

Query Match 47.6%; Score 379; DB 10; Length 352;
Best Local Similarity 35.0%; Pred. No. 1.4e-29;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

QY 1 MDYQVSSPIVDINYYTSEPCQKINVKQIAA----- 30
DB 1 MDYQVSSPIVDINYYTSEPCQKINVKQIAALLPPLYSLVFIFFGVGNMLVILINCKR 60
QY 31 -----YKGLC-----AAQWDFGNTMCOHQRVHGHHSYKCG--- 64
DB 61 LKSMTDIYLLNLAISDLFLLTVFPFAHYAAAQWDFGNTMC--QLLTGLYFIFFSGIFF 118
QY 65 -----YKGLC-----LCTRSQKEGLHYTC 78
DB 119 ILLTIDRYLAVHVAFAKARTVTFGVTSVITWVAVFASLPGLIFTRSQKEGLHYTC 178
QY 79 SSHFPYSQYQFKNFOTLKI-----HORVHGG----- 105
DB 179 SSHFPYSQYQFKNFOTLKIIVILGLVPLLMVICYSGLKTLRLCRNEKKHRAVRLIF 238
QY 106 -----GGSYKGLC-----QEFFGLNCCSSNRLDGHQRV 135
DB 239 TIMIVYFLFAPYNIIVLLNTFOEFFGLNCCSSNRLDQAMQV 281

RESULT 15
US-09-502-783A-2
Sequence 2, Application US/09502783A
Patent No. US20020132269A1
GENERAL INFORMATION:
APPLICANT: Li, Yi
APPLICANT: Ruben, Steven M.
TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein Chemokine Receptor (CCRS)
FILE REFERENCE: HDGNR10
FILE REFERENCE: 1488.1150006
CURRENT APPLICATION NUMBER: US/09/502,783A
CURRENT FILING DATE: 2001-08-23
PRIOR APPLICATION NUMBER: 08/466,343
PRIOR FILING DATE: 1995-06-06
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 352
TYPE: PRT
ORGANISM: Homo sapiens
US-09-502-783A-2

Query Match 47.6%; Score 379; DB 10; Length 352;
Best Local Similarity 35.0%; Pred. No. 1.4e-29;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

```
QY 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 30
Db |||||
QY 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPLYSLVFI FGVGNMLVILINCKR 60
Db |||||
QY 31 -----YKGLC----- 64
Db |||||
QY 61 LKSMTDIYLLNLAISDLFFLLTVPFWAHYAAQWDFGNTMC--QLLTGLYFIFGFFSGIFF 118
Db |||||
QY 65 -----LCTRSQKEGLHYTC 78
Db |||||
QY 119 IILLTIDRYLAVVHAVPALKARTVTFGVTSVITWVAVFASLPGLIFTRSQKEGLHYTC 178
Db |||||
QY 79 SSHFPYSQYQFWKNFOTLKI-----HORVHGG----- 105
Db |||||
QY 179 SSHFPYSQYQFWKNFOTLKIIVILGLVPLIVMVICYSGILKTLRCRNEKGRHRAVLIF 238
Db |||||
QY 106 -----GGSYKGLC-----OEFGLNCCSSNRLDGHQV 135
Db |||||
QY 239 TIMIVYFLFWAPYIVLLNTFOEFFGLNCCSSNRLDQAMQV 281
```

Search completed: October 21, 2003, 18:07:21
Job time : 44.1017 secs

GerCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:27:21 ; Search time 16.6667 Seconds
(without alignments)
15.232 Million cell updates/sec

Title: US-10-057-890A-15

Perfect score: 6

Sequence: 1 GHHHS 6

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 90058

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Listing first 150 summaries

Database : Issued Patents AA.*

1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/PCTUS.COMB.pep.*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5	83.3	7	1	US-08-100-744-7
2	5	83.3	7	1	US-08-284-784-7
3	5	83.3	7	2	US-08-854-811-7
4	5	83.3	9	1	US-08-155-171B-4
5	5	83.3	9	1	US-08-323-445A-2
6	5	83.3	9	1	US-08-515-903A-2
7	5	83.3	9	2	US-08-435-998-4
8	5	83.3	9	5	PCT-US95-12840-2
9	5	83.3	10	1	US-07-807-539A-73
10	5	83.3	10	2	US-08-482-142-150
11	5	83.3	10	2	US-08-478-572-150
12	5	83.3	10	3	US-08-300-928C-88
13	5	83.3	10	3	US-08-430-944D-88
14	5	83.3	10	3	US-08-430-014-88
15	5	83.3	10	3	US-08-431-184-88
16	5	83.3	10	3	US-08-163-919A-17
17	5	83.3	10	3	US-08-556-978B-10
18	5	83.3	10	3	US-08-484-296-150
19	5	83.3	10	5	PCT-US94-14073-17
20	4	66.7	4	2	US-08-623-833B-57
21	4	66.7	4	4	US-08-461-713-41
22	4	66.7	5	1	US-07-826-928A-3
23	4	66.7	5	1	US-07-826-928A-7
24	4	66.7	5	1	US-08-480-604A-32
25	4	66.7	5	1	US-08-651-323A-12
26	4	66.7	5	2	US-08-623-833B-55
27	4	66.7	5	2	US-08-405-496A-30

28	4	66.7	5	3	US-08-915-136-32	Sequence 32, Appl
29	4	66.7	6	1	US-08-160-670A-6	Sequence 6, Appl
30	4	66.7	6	1	US-08-140-222-1	Sequence 1, Appl
31	4	66.7	6	1	US-08-131-365B-48	Sequence 48, Appl
32	4	66.7	6	1	US-08-469-486-47	Sequence 47, Appl
33	4	66.7	6	1	US-08-460-343B-51	Sequence 51, Appl
34	4	66.7	6	1	US-08-299-567-1	Sequence 1, Appl
35	4	66.7	6	1	US-08-398-028B-51	Sequence 51, Appl
36	4	66.7	6	1	US-08-660-626-4	Sequence 4, Appl
37	4	66.7	6	1	US-08-434-705B-17	Sequence 17, Appl
38	4	66.7	6	1	US-08-470-837-22	Sequence 22, Appl
39	4	66.7	6	2	US-08-472-244-14	Sequence 14, Appl
40	4	66.7	6	2	US-08-504-265B-51	Sequence 51, Appl
41	4	66.7	6	2	US-08-623-833B-16	Sequence 16, Appl
42	4	66.7	6	2	US-08-652-507-8	Sequence 8, Appl
43	4	66.7	6	2	US-08-668-123-48	Sequence 48, Appl
44	4	66.7	6	2	US-08-811-028-49	Sequence 49, Appl
45	4	66.7	6	2	US-08-469-658-47	Sequence 47, Appl
46	4	66.7	6	2	US-09-086-201-17	Sequence 17, Appl
47	4	66.7	6	2	US-08-591-196-57	Sequence 57, Appl
48	4	66.7	6	2	US-08-596-139-18	Sequence 18, Appl
49	4	66.7	6	3	US-08-481-435-42	Sequence 42, Appl
50	4	66.7	6	3	US-08-772-440-28	Sequence 28, Appl
51	4	66.7	6	3	US-09-267-031-18	Sequence 18, Appl
52	4	66.7	6	3	US-09-398-341-1	Sequence 1, Appl
53	4	66.7	6	3	US-09-031-168-4	Sequence 4, Appl
54	4	66.7	6	3	US-08-789-133B-99	Sequence 99, Appl
55	4	66.7	6	3	US-08-995-659-18	Sequence 18, Appl
56	4	66.7	6	3	US-09-142-334-26	Sequence 26, Appl
57	4	66.7	6	3	US-09-215-649A-18	Sequence 18, Appl
58	4	66.7	6	4	US-09-140-084-7	Sequence 7, Appl
59	4	66.7	6	4	US-09-382-950-4	Sequence 4, Appl
60	4	66.7	6	4	US-09-382-736B-5	Sequence 5, Appl
61	4	66.7	6	4	US-08-977-378-2	Sequence 2, Appl
62	4	66.7	6	4	US-08-353-555-2	Sequence 2, Appl
63	4	66.7	6	4	US-08-868-452-22	Sequence 22, Appl
64	4	66.7	6	4	US-09-000-094-47	Sequence 47, Appl
65	4	66.7	6	4	US-09-058-483-7	Sequence 7, Appl
66	4	66.7	6	4	US-09-140-201-8	Sequence 8, Appl
67	4	66.7	6	4	US-09-039-780A-106	Sequence 106, App
68	4	66.7	6	4	US-09-291-170A-14	Sequence 14, Appl
69	4	66.7	6	4	US-09-577-780-18	Sequence 18, Appl
70	4	66.7	6	4	US-09-724-297-7	Sequence 7, Appl
71	4	66.7	6	4	US-09-367-309A-4	Sequence 4, Appl
72	4	66.7	6	4	US-09-724-884-14	Sequence 14, Appl
73	4	66.7	6	4	US-09-529-279-6	Sequence 6, Appl
74	4	66.7	6	4	US-09-313-942-1	Sequence 1, Appl
75	4	66.7	6	4	US-09-353-215-6	Sequence 6, Appl
76	4	66.7	6	4	US-09-577-800-18	Sequence 18, Appl
77	4	66.7	6	4	US-09-667-422-7	Sequence 7, Appl
78	4	66.7	6	4	US-09-315-926A-81	Sequence 81, Appl
79	4	66.7	6	4	US-09-731-558-11	Sequence 11, Appl
80	4	66.7	6	4	US-09-280-030-61	Sequence 61, Appl
81	4	66.7	6	4	US-09-466-496-18	Sequence 18, Appl
82	4	66.7	6	4	US-09-648-569A-39	Sequence 39, Appl
83	4	66.7	6	4	US-09-871-856-18	Sequence 18, Appl
84	4	66.7	6	4	US-10-158-895-6	Sequence 6, Appl
85	4	66.7	6	4	US-09-904-196B-9	Sequence 9, Appl
86	4	66.7	6	4	US-09-871-291-18	Sequence 18, Appl
87	4	66.7	6	4	US-09-115-397-1	Sequence 1, Appl
88	4	66.7	6	4	US-09-463-194-2	Sequence 2, Appl
89	4	66.7	6	5	PCT-US94-04361-55	Sequence 55, Appl
90	4	66.7	7	1	US-08-276-852-14	Sequence 14, Appl
91	4	66.7	7	1	US-08-899-575-14	Sequence 14, Appl
92	4	66.7	7	1	US-08-899-575-14	Sequence 14, Appl
93	4	66.7	7	3	US-08-865-468-9	Sequence 9, Appl
94	4	66.7	7	3	PCT-US95-08743-14	Sequence 14, Appl
95	4	66.7	8	2	US-08-762-433-6	Sequence 6, Appl
96	4	66.7	8	2	US-09-001-219-6	Sequence 6, Appl
97	4	66.7	8	2	US-08-481-658B-22	Sequence 22, Appl
98	4	66.7	8	2	US-08-477-504A-22	Sequence 22, Appl
99	4	66.7	8	2	US-08-486-756A-22	Sequence 22, Appl
100	4	66.7	8	2		

101 Sequence 22, Appl
102 Sequence 22, Appl
103 Sequence 22, Appl
104 Sequence 11, Appl
105 Sequence 22, Appl
106 Sequence 22, Appl
107 Sequence 57, Appl
108 Sequence 13, Appl
109 Sequence 40, Appl
110 Sequence 31, Appl
111 Sequence 22, Appl
112 Sequence 22, Appl
113 Sequence 15, Appl
114 Sequence 16, Appl
115 Sequence 14, Appl
116 Sequence 40, Appl
117 Sequence 10, Appl
118 Sequence 12, Appl
119 Sequence 12, Appl
120 Sequence 12, Appl
121 Sequence 12, Appl
122 Sequence 19, Appl
123 Sequence 1, Appl
124 Sequence 9, Appl
125 Sequence 36, Appl
126 Sequence 102, Appl
127 Sequence 1, Appl
128 Sequence 17, Appl
129 Sequence 8, Appl
130 Sequence 16, Appl
131 Sequence 5, Appl
132 Sequence 17, Appl
133 Sequence 16, Appl
134 Sequence 10, Appl
135 Sequence 14, Appl
136 Sequence 30, Appl
137 Sequence 31, Appl
138 Sequence 16, Appl
139 Sequence 13, Appl
140 Sequence 3, Appl
141 Sequence 24, Appl
142 Sequence 14, Appl
143 Sequence 14, Appl
144 Sequence 9, Appl
145 Sequence 17, Appl
146 Sequence 129, Appl
147 Sequence 129, Appl
148 Sequence 101, Appl
149 Sequence 27, Appl
150 Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-08-100-744-7
; Sequence 7, Application US/08100744
; Patent No. 5629172
; GENERAL INFORMATION:
; APPLICANT: MASCARENHAS, DESMOND
; APPLICANT: ZHANG, SUNNY
; APPLICANT: OLSON, PAMELA
; APPLICANT: OLSEN, DAVID
; APPLICANT: CARRILLO, PEDRO A.
; TITLE OF INVENTION: POLYPEPTIDE FUSIONS TO
; TITLE OF INVENTION: INTERLEUKIN-1-LIKE POLYPEPTIDES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California

; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/100,744
; FILING DATE: 02-AUG-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: LUTHER, BARBARA J.
; REGISTRATION NUMBER: 33,954
; REFERENCE/DOCKET NUMBER: 22095-20275.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-100-744-7

Query Match 83.3%; Score 5; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e-05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 HHHS 6
Db 3 HHHS 7

RESULT 2
US-08-284-784-7
; Sequence 7, Application US/08284784
; Patent No. 5629172
; GENERAL INFORMATION:
; APPLICANT: MASCARENHAS, DESMOND
; APPLICANT: ZHANG, YANG
; APPLICANT: OLSON, PAMELA S.
; APPLICANT: OLSEN, DAVID R.
; APPLICANT: CARRILLO, PEDRO A.
; TITLE OF INVENTION: EXPRESSION OF FUSION POLYPEPTIDES
; TITLE OF INVENTION: TRANSPORTED OUT OF THE CYTOPLASM WITHOUT LEADER SEQUENCES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/284,784
; FILING DATE: 02-AUG-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: PARK, FREDDIE K.
; REGISTRATION NUMBER: 35,636
; REFERENCE/DOCKET NUMBER: 22095-20275.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141

; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-284-784-7

Query Match 83.3%; Score 5; DB 1; Length 7;
Best Local Similarity 100.0%; Pred.No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHS 6
DB 3 HHHS 7

RESULT 3
US-08-854-811-7
; Sequence 7, Application US/08854811
; Patent No. 5914254
; GENERAL INFORMATION:
; APPLICANT: Mascarenhas, Desmond
; APPLICANT: Zhang, Yang
; APPLICANT: Olson, Pamela S.
; APPLICANT: Olsen, David R.
; APPLICANT: Cohen, Pedro A.
; TITLE OF INVENTION: EXPRESSION OF FUSION POLYPEPTIDES
; TITLE OF INVENTION: TRANSPORTED OUT OF THE CYTOPLASM WITHOUT LEADER
; TITLE OF INVENTION: SEQUENCES
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,811
; FILING DATE: 12-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/284,784
; FILING DATE: 02-AUG-1994
; APPLICATION NUMBER: 08/100,744
; FILING DATE: 02-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Ruffinger, Nicholas S
; REGISTRATION NUMBER: 39,124
; REFERENCE/DOCKET NUMBER: 22095-20275.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-854-811-7

Query Match 83.3%; Score 5; DB 2; Length 7;
Best Local Similarity 100.0%; Pred.No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHS 6

DB 3 HHHS 7

RESULT 4
US-08-155-171B-4
; Sequence 4, Application US/08155171B
; Patent No. 5543264
; GENERAL INFORMATION:
; APPLICANT: Anderson, Carl W.
; APPLICANT: Mangel, Walter F.
; TITLE OF INVENTION: Co-Factor Activated Recombinant
; TITLE OF INVENTION: Adenovirus Proteinases (As Amended)
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/155,171B
; FILING DATE: 19-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/851,217
; FILING DATE: 13-MAR-1992

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/545,585
; FILING DATE: 29-JUN-1990

ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: BNL91-01A2, AUI93-22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-155-171B-4

Query Match 83.3%; Score 5; DB 1; Length 9;
Best Local Similarity 100.0%; Pred.No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHS 5
DB 1 GHHS 5

RESULT 5
US-08-323-445A-2
; Sequence 2, Application US/08323445A
; Patent No. 5763733
; GENERAL INFORMATION:
; APPLICANT: Whitlow, Marc
; APPLICANT: Bilpula, David
; APPLICANT: Shorri, Robert
; TITLE OF INVENTION: Antigen-Binding Fusion Proteins
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600

CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/323.445A
 FILING DATE: 13-Oct-1994
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: Bugalsky, Lawrence B.
 REGISTRATION NUMBER: 35,086
 REFERENCE/DOCKET NUMBER: 0977.2060000
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2600
 TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 9 amino acids
 TYPE: amino acid
 TOPOLOGY: both
 US-08-323-445A-2

Query Match 83.3%; Score 5; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHS 6
 DB 4 HHHS 8

RESULT 6
 US-08-515-903A-2
 Sequence 2, Application US/08515903A
 Patent No. 5767260
 GENERAL INFORMATION:
 APPLICANT: Whitlow, Marc
 APPLICANT: Filipula, David
 APPLICANT: Short, Robert
 TITLE OF INVENTION: Antigen-Binding Fusion Proteins
 NUMBER OF SEQUENCES: 19
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.
 STREET: 1100 New York Avenue, Suite 600
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20005-3934
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/515,903A
 FILING DATE: 16-AUG-1995
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: Goldstein, Jorge A.
 REGISTRATION NUMBER: 29,021
 REFERENCE/DOCKET NUMBER: 0977.2060001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2600
 TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 9 amino acids
 TYPE: amino acid

TOPOLOGY: both
 US-08-515-903A-2
 Query Match 83.3%; Score 5; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHS 6
 DB 4 HHHS 8

RESULT 7
 US-08-435-998-4
 Sequence 4, Application US/08435998
 Patent No. 5935840
 GENERAL INFORMATION:
 APPLICANT: Anderson, Carl W.
 APPLICANT: Mangel, Walter F.
 TITLE OF INVENTION: Co-Factor Activated Recombinant
 TITLE OF INVENTION: Adenovirus Proteinases (As Amended)
 NUMBER OF SEQUENCES: 45
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
 STREET: Two Militia Drive
 CITY: Lexington
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02173
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/435,998
 FILING DATE: 05-MAY-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 06/155,171
 FILING DATE: 19-NOV-1993
 APPLICATION NUMBER: US 07/851,217
 FILING DATE: 13-MAR-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/545,585
 FILING DATE: 29-JUN-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Granahan, Patricia
 REGISTRATION NUMBER: 32,227
 REFERENCE/DOCKET NUMBER: BNL91-01A2, AU193-22
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 861-6240
 TELEFAX: (617) 861-9540
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 9 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 US-08-435-998-4

Query Match 83.3%; Score 5; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHH 5
 DB 1 GHHH 5

RESULT 8
 PCT-US95-12840-2
 Sequence 2, Application PC/TUS9512840

;
; GENERAL INFORMATION:
; APPLICANT: ENZON, INC.
; TITLE OF INVENTION: Antigen-Binding Fusion Proteins
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/12840
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: US 08/323,445
; FILING DATE: 13-OCT-1994
; CLASSIFICATION:
; NAME: Goldstein, Jorge A.
; REGISTRATION NUMBER: 29,021
; REFERENCE/DOCKET NUMBER: 0977.206PC00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; PCT-US95-12840-2

Query Match 83.3%; Score 5; DB 5; Length 9;
Best Local Similarity 100.0%; Pred.No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHS 6
DB 4 HHHS 8

RESULT 9
US-07-807-529A-73
; Sequence 73, Application US/07807529A
; Patent No. 5547869
; GENERAL INFORMATION:
; APPLICANT: Rogers, Bruce L.
; APPLICANT: Morgenstern, Jay
; APPLICANT: Bond, Julian F.
; APPLICANT: Garman, Richard D.
; APPLICANT: Greenstein, Julia L.
; APPLICANT: Kuo, Mei-chang
; APPLICANT: Morville, Malcolm
; TITLE OF INVENTION: RECOMBITOPE PEPTIDES
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
; STREET: One Kendall Square, Building 600
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02139
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/807,529A
; FILING DATE: 19911213
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/662,276
; FILING DATE: 28-FEB-1991
; APPLICATION NUMBER: US 07/431,565
; FILING DATE: 03-NOV-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Channing, Stacey L.
; REGISTRATION NUMBER: 31,095
; REFERENCE/DOCKET NUMBER: IPC-027/imi-015
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 494-0060
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-07-807-529A-73

Query Match 83.3%; Score 5; DB 1; Length 10;
Best Local Similarity 100.0%; Pred.No. 0.16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHH 5
DB 2 GHHH 6

RESULT 10
US-08-482-142-150
; Sequence 150, Application US/08482142
; Patent No. 5820862
; GENERAL INFORMATION:
; APPLICANT: Garman, Richard
; APPLICANT: Greenstein, Julia
; APPLICANT: Kuo, Mei-chang
; APPLICANT: Rogers, Bruce
; APPLICANT: Franzen, Henry
; APPLICANT: Chen, Xian
; APPLICANT: Evans, Sean
; APPLICANT: Shaked, Ze'ev
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
; FROM DERMATOPHAGOIDES (HOUSE DUST MITE)
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,142
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,307
; FILING DATE: 07 June 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG, ANNE I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: 017.6US
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 466-6000
TELEFAX: (617) 466-6040
INFORMATION FOR SEQ ID NO: 150:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-482-142-150

Query Match 83.3%; Score 5; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHHH 5
DB 2 GHHHH 6

RESULT 11

US-08-478-572-150
Sequence 150, Application US/08478572
Patent No. 5968526

GENERAL INFORMATION:

APPLICANT: Garman, Richard
APPLICANT: Greenstein, Julia
APPLICANT: Kuo, Mei-chang
APPLICANT: Rogers, Bruce
APPLICANT: Franzen, Henry
APPLICANT: Chen, Xian
APPLICANT: Evans, Sean
APPLICANT: Shaked, Ze'ev
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)
NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
STREET: 610 LINCOLN STREET
CITY: WALTHAM
STATE: MA
COUNTRY: USA
ZIP: 02154

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/478,572
FILING DATE: 07-June-1995
CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/445,307

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: CRAIG, ANNE I.

REGISTRATION NUMBER: 32,976

REFERENCE/DOCKET NUMBER: 017.6US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 466-6000

TELEFAX: (617) 466-6040

INFORMATION FOR SEQ ID NO: 150:

SEQUENCE CHARACTERISTICS:

LENGTH: 10 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: internal

US-08-478-572-150

Query Match 83.3%; Score 5; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.16;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GHHHH 5
DB 2 GHHHH 6

RESULT 12

US-08-300-928C-88
Sequence 88, Application US/08300928C
Patent No. 6019972

GENERAL INFORMATION:
APPLICANT: GEFTER, Malcolm L. et al.
TITLE OF INVENTION: PEPTIDES FOR HUMAN T CELL REACTIVE FELINE
TITLE OF INVENTION: PROTEIN (TRFP)
NUMBER OF SEQUENCES: 101
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
STREET: 610 LINCOLN STREET
CITY: WALTHAM
STATE: Massachusetts
COUNTRY: USA
ZIP: 02145

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/300,928C
FILING DATE: September 2, 1994
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/807,529

FILING DATE: December 13, 1991

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: AMY E. MANDRAGOURAS

REGISTRATION NUMBER: 36,207

REFERENCE/DOCKET NUMBER: 002.6US (IMI-044)

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 88:

SEQUENCE CHARACTERISTICS:

LENGTH: 10 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: internal

US-08-300-928C-88

Query Match 83.3%; Score 5; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.16;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHHH 5
DB 2 GHHHH 6

RESULT 13

US-08-430-944D-88

Sequence 88, Application US/08430944D

Patent No. 6025162

GENERAL INFORMATION:

APPLICANT: Bruce L. Rogers et al.

TITLE OF INVENTION: A HUMAN T CELL REACTIVE FELINE PROTEIN

TITLE OF INVENTION:

NUMBER OF SEQUENCES: 103

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD, LLP

STREET: 28 State Street

```
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/430.944D
; FILING DATE: 28-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/430.014
; FILING DATE: 27-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/300.928
; FILING DATE: 02-SEPT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Amy E. Mandragouras
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: IMI-044DV2
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 88:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-08-430-944D-88

Query Match      83.3%; Score 5; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GHHHH 5
Db      2 GHHHH 6

RESULT 14
US-08-430-014-88
; Sequence 88, Application US/08430014
; Patent No. 6048962
; GENERAL INFORMATION:
; APPLICANT: GEFTER, Malcolm L. et al.
; TITLE OF INVENTION: PEPTIDES FOR HUMAN T CELL REACTIVE FELINE
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMUNOLOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02145
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/430.014
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/300.928
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: AMY E. MANDRAGOURAS
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: 002.6US (IMI-044)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 88:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-08-430-014-88

Query Match      83.3%; Score 5; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GHHHH 5
Db      2 GHHHH 6

RESULT 15
US-08-431-184-88
; Sequence 88, Application US/08431184
; Patent No. 6120769
; GENERAL INFORMATION:
; APPLICANT: Bruce L. Rogers et al.
; TITLE OF INVENTION: A HUMAN T CELL REACTIVE FELINE PROTEIN
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/431.184
; FILING DATE: 28-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/430.014
; FILING DATE: 27-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/300.928
; FILING DATE: 02-SEPT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Amy E. Mandragouras
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: IMI-044DV3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 88:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-08-431-184-88

Query Match      83.3%; Score 5; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 GHHH 5
 Db 2 GHHH 6

RESULT 16

US-08-163-919A-17
 ; Sequence 17, Application US/08163919A
 ; Patent No. 6180771
 ; GENERAL INFORMATION:
 ; APPLICANT: THOMAS, Wayne R., CHUA, Kaw-Yan, ROGERS, Bruce L., and
 ; APPLICANT: KUO, Mei-chang
 ; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING A HOUSE DUST MITE
 ; TITLE OF INVENTION: ALLERGEN, DER P III, AND USES THEREFOR
 ; NUMBER OF SEQUENCES: 18
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Lahive & Cockfield
 ; STREET: 60 State Street suite 510
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII TEXT
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/163,919A
 ; FILING DATE: 08-DEC-1993
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mandragoras, Amy E.
 ; REGISTRATION NUMBER: 36,207
 ; REFERENCE/DOCKET NUMBER: IPC-072 (IMI-041)
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 227-7400
 ; TELEFAX: (617) 227-5941
 ; INFORMATION FOR SEQ ID NO: 17:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 10 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; FRAGMENT TYPE: internal
 ; US-08-163-919A-17

Query Match 83.3%; Score 5; DB 3; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.16;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHH 5
 Db 2 GHHH 6

RESULT 17

US-08-556-978B-10
 ; Sequence 10, Application US/08556978B
 ; Patent No. 6268169
 ; GENERAL INFORMATION:
 ; APPLICANT: FARNESSTOCK, STEPHEN F.
 ; TITLE OF INVENTION: NOVEL RECOMBINANTLY PRODUCED
 ; TITLE OF INVENTION: SPIDER SILK ANALOGS
 ; NUMBER OF SEQUENCES: 107
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
 ; STREET: 1007 MARKET STREET
 ; CITY: WILMINTON

STATE: DELAWARE
 COUNTRY: UNITED STATES OF AMERICA
 ZIP: 19898
 COMPUTER READABLE FORM:
 ; MEDIUM TYPE: DISKETTE, 3.50 INCH
 ; COMPUTER: IBM PC COMPATIBLE
 ; OPERATING SYSTEM: MICROSOFT WINDOWS 95
 ; SOFTWARE: MICROSOFT WORD FOR WINDOWS 95
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/556,978B
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION NUMBER:
 ; FILING DATE: JUNE 15, 1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: FLOYD, LINDA AXAMETHY
 ; REGISTRATION NUMBER: 33,692
 ; REFERENCE/DOCKET NUMBER: CR-9389-A
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 302-892-8112
 ; TELEFAX: 302-773-0164
 ; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 10 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: unknown
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: peptide
 ; US-08-556-978B-10

Query Match 83.3%; Score 5; DB 3; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.16;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHS 6
 Db 5 HHHS 9

RESULT 18

US-08-484-296-150
 ; Sequence 150, Application US/08484296
 ; Patent No. 6268491
 ; GENERAL INFORMATION:
 ; APPLICANT: Garman, Richard
 ; APPLICANT: Greenstein, Julia
 ; APPLICANT: Kuo, Mei-chang
 ; APPLICANT: Rogers, Bruce
 ; APPLICANT: Franzen, Henry
 ; APPLICANT: Chen, Xian
 ; APPLICANT: Evans, Sean
 ; APPLICANT: Shaked, Ze'ev
 ; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
 ; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)
 ; NUMBER OF SEQUENCES: 207
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
 ; STREET: 610 LINCOLN STREET
 ; CITY: WALTHAM
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02154
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII TEXT
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/484,296
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/445,307
FILING DATE: 07 June 1995
ATTORNEY/AGENT INFORMATION:
NAME: CRAIG, ANNE I.
REGISTRATION NUMBER: 32,976
REFERENCE/DOCKET NUMBER: 017.6US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 466-6000
TELEFAX: (617) 466-6040
INFORMATION FOR SEQ ID NO: 150:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-484-296-150

Query Match 83.3%; Score 5; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHHH 5
DB 2 GHHHH 6

RESULT 19
PCT-US94-14073-17
Sequence 17, Application PC/TUS9414073
GENERAL INFORMATION:
APPLICANT: THOMAS, Wayne R., CHUA, Kaw-Yan, ROGERS, Bruce L., and
APPLICANT: KUO, Mel-chang
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING A HOUSE DUST MITE
TITLE OF INVENTION: ALLERGEN, DER P III, AND USES THEREFOR
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lahive & Cockfield
STREET: 60 State Street suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/14073
FILING DATE: 08-DEC-1993
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IPC-072 (IMI-041)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
PCT-US94-14073-17

Query Match 83.3%; Score 5; DB 5; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHHH 5
DB 2 GHHHH 6

RESULT 20
US-08-623-833B-57
Sequence 57, Application US/08623833B
Patent No. 5866683
GENERAL INFORMATION:
APPLICANT: SHIMURA, Kiyohito
APPLICANT: KASAI, Kenichi
APPLICANT: MATSUMOTO, Hiroyuki
APPLICANT: TAKAMOTO, Hisayoshi
TITLE OF INVENTION: ISOELECTRIC POINT MARKERS FOR
TITLE OF INVENTION: ISOELECTRIC FOCUSING WITH FLUORESCENCE DETECTION
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/623,833B
FILING DATE: 29-MAR-1996
CLASSIFICATION: S30
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 076863/1995
FILING DATE: 31-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 271196/1995
FILING DATE: 19-OCT-1995
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acid
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-623-833B-57

Query Match 66.7%; Score 4; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 1 HHHH 4

RESULT 21
US-08-840-713-41
Sequence 41, Application US/08840713
Patent No. 6498233
GENERAL INFORMATION:
APPLICANT: WELS, Winfried, Dr.
APPLICANT: FOYMINAYA, Jesus
TITLE OF INVENTION: NUCLEIC ACID TRANSFER SYSTEM
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
STREET: 655 15th St., N.W., Suite 330 - G St. Lobby
CITY: Washington
STATE: D.C.

COUNTRY: USA
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/840,713
FILING DATE: 25-APR-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kitts, Monica Chin
REGISTRATION NUMBER: 36,105
REFERENCE/DOCKET NUMBER: 1614-7014
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 638 - 5000
TELEFAX: (202) 638 - 4810
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-840-713-41

Query Match 66.7%; Score 4; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
Db 1 HHHH 4

RESULT 22
US-07-826-928A-3
Sequence 3, Application US/07826928A
Patent No. 5439829
GENERAL INFORMATION:
APPLICANT: Anderson, Leslie D.
APPLICANT: Cook, James A.
APPLICANT: David, Gary S.
APPLICANT: Hochschwender, Susan M.
APPLICANT: Kasher, Mary S.
APPLICANT: Smith, Michele C.
APPLICANT: Stemmer, Willem P.
TITLE OF INVENTION: METHOD OF IMMOBILIZING AND CROSS LINKING
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/826,928A
FILING DATE: 19920124
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy, Richard B.
REGISTRATION NUMBER: 35,296
REFERENCE/DOCKET NUMBER: X8180A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3589
TELEFAX: 317-276-1294

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-826-928A-3

Query Match 66.7%; Score 4; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
Db 1 HHHH 4

RESULT 23
US-07-826-928A-7
Sequence 7, Application US/07826928A
Patent No. 5439829
GENERAL INFORMATION:
APPLICANT: Anderson, Leslie D.
APPLICANT: Cook, James A.
APPLICANT: David, Gary S.
APPLICANT: Hochschwender, Susan M.
APPLICANT: Kasher, Mary S.
APPLICANT: Smith, Michele C.
APPLICANT: Stemmer, Willem P.
TITLE OF INVENTION: METHOD OF IMMOBILIZING AND CROSS LINKING
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/826,928A
FILING DATE: 19920124
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy, Richard B.
REGISTRATION NUMBER: 35,296
REFERENCE/DOCKET NUMBER: X8180A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3589
TELEFAX: 317-276-1294
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-826-928A-7

Query Match 66.7%; Score 4; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
Db 1 HHHH 4

RESULT 24
US-08-480-604A-32
; Sequence 32, Application US/08480604A
; Patent No. 5736139
; GENERAL INFORMATION:
; APPLICANT: KINK, JOHN A.
; APPLICANT: THALLEY, BRUCE S.
; APPLICANT: PADHYE, NISHA V.
; APPLICANT: FIRCA, JOSEPH R.
; APPLICANT: STAFFORD, DOUGLAS C.
; TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
; TITLE OF INVENTION: PREVENTION OF C. DIFFICILE DISEASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480.604A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/422,711
; FILING DATE: 14-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/405,496
; FILING DATE: 16-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/329,154
; FILING DATE: 25-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/161,907
; FILING DATE: 02-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/985,321
; FILING DATE: 04-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/429,791
; FILING DATE: 31-OCT-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: INGOLIA, DIANE E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: OPHD-01763
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-480-604A-32

Query Match 66.7%; Score 4; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 1 HHHH 4

RESULT 25
US-08-651-323A-12
; Sequence 12, Application US/08651323A
; Patent No. 5798226
; GENERAL INFORMATION:
; APPLICANT: Lehle, Ludwig; Kopetzki, Erhard
; TITLE OF INVENTION: Yeast Host Strains With Defects in
; TITLE OF INVENTION: N-Glycosylation
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.25 inch, 1.44mb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/651,323A
; FILING DATE: May 31, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/104,436
; FILING DATE: August 9, 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 42 26 094.9
; FILING DATE: 07-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 43 01 932.3
; FILING DATE: 23-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Andrew L. Tiajolloff
; REGISTRATION NUMBER: 31,575
; REFERENCE/DOCKET NUMBER: BOER-1018.2 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-651-323A-12

Query Match 66.7%; Score 4; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 2 HHHH 5

RESULT 26
US-08-623-833B-55
; Sequence 55, Application US/08623833B
; Patent No. 5866883
; GENERAL INFORMATION:
; APPLICANT: SHIMURA, Kiyohito
; APPLICANT: KASAI, Kenichi
; APPLICANT: MATSUMOTO, Hiroyuki
; APPLICANT: TAKAMOTO, Hisayoshi
; TITLE OF INVENTION: ISOELECTRIC POINT MARKERS FOR
; TITLE OF INVENTION: ISOELECTRIC FOCUSING WITH FLUORESCENCE DETECTION
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.
; STREET: 1100 New York Avenue, N.W.

CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20005-3918
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WordPerfect 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/623,833B
 FILING DATE: 29-MAR-1996
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 076863/1995
 FILING DATE: 31-MAR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 271196/1995
 FILING DATE: 19-OCT-1995
 INFORMATION FOR SEQ ID NO: 55:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5 amino acid
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-623-833B-55

Query Match 66.7%; Score 4; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
 DB 1 HHHH 4

RESULT 27
 US-08-405-496A-30
 Sequence 30, Application US/08405496A
 Patent No. 5919665
 GENERAL INFORMATION:
 APPLICANT: WILLIAMS, JAMES A.
 TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM
 TITLE OF INVENTION: NEUROTOXIN
 NUMBER OF SEQUENCES: 30
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MEDLEN & CARROLL, LLP
 STREET: 220 MONTGOMERY STREET, SUITE 2200
 CITY: SAN FRANCISCO
 STATE: CALIFORNIA
 COUNTRY: USA
 ZIP: 94104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/405,496A
 FILING DATE: 16-MAR-1995
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/329,154
 FILING DATE: 25-OCT-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/161,907
 FILING DATE: 02-DEC-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/985,321
 FILING DATE: 04-DEC-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/429,791
 FILING DATE: 31-OCT-1989

ATTORNEY/AGENT INFORMATION:
 NAME: INGOLIA, DIANE E.
 REGISTRATION NUMBER: 40,027
 REFERENCE/DOCKET NUMBER: OPHD-01308
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 705-8410
 TELEFAX: (415) 397-8338
 INFORMATION FOR SEQ ID NO: 30:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: not relevant
 MOLECULE TYPE: peptide
 US-08-405-496A-30
 Query Match 66.7%; Score 4; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
 DB 1 HHHH 4

RESULT 28
 US-08-915-136-32
 Sequence 32, Application US/08915136
 Patent No. 6290960
 GENERAL INFORMATION:
 APPLICANT: KINK, JOHN A.
 APPLICANT: THALLEY, BRUCE S.
 APPLICANT: PADHYE, NISHA V.
 APPLICANT: FIRCA, JOSEPH R.
 APPLICANT: STAFFORD, DOUGLAS C.
 TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
 TITLE OF INVENTION: PREVENTION OF C. DIFFICILE DISEASE
 NUMBER OF SEQUENCES: 32
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MEDLEN & CARROLL, LLP
 STREET: 220 MONTGOMERY STREET, SUITE 2200
 CITY: SAN FRANCISCO
 STATE: CALIFORNIA
 COUNTRY: UNITED STATES OF AMERICA
 ZIP: 94104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/915,136
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA: 08/480,604
 FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/405,496
 FILING DATE: 16-MAR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/329,154
 FILING DATE: 25-OCT-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/161,907
 FILING DATE: 02-DEC-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/985,321
 FILING DATE: 04-DEC-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/429,791
 FILING DATE: 31-OCT-1989
 ATTORNEY/AGENT INFORMATION:

NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPND-01763
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-915-136-32

Query Match 66.7%; Score 4; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
Db 1 HHHH 4

RESULT 29
US-08-160-670A-6
Sequence 6, Application US/08160670A
Patent No. 5449758
GENERAL INFORMATION:
APPLICANT: Hartley, James L.
TITLE OF INVENTION: Protein Size Marker Ladder
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/160.670A
FILING DATE: 12/2/93
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.2580000
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: both
US-08-160-670A-6

Query Match 66.7%; Score 4; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
Db 1 HHHH 4

RESULT 30
US-08-140-222-1

Sequence 1, Application US/08140222
Patent No. 5470952
GENERAL INFORMATION:
APPLICANT: Stahl, et al.
TITLE OF INVENTION: CNTF Family Antagonists
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Regeneron Pharmaceuticals, Inc.
STREET: 777 Old Saw Mill River Road
CITY: Tarrytown
STATE: New York
COUNTRY: U.S.A.
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/140.222
FILING DATE: 20-OCT-1993
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kempler Ph.D., Gail M.
REGISTRATION NUMBER: 32,143
REFERENCE/DOCKET NUMBER: REG 200
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-347-7000
TELEFAX: 914-347-2113
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-140-222-1

Query Match 66.7%; Score 4; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
Db 1 HHHH 4

RESULT 31
US-08-131-365B-48
Sequence 48, Application US/08131365B
Patent No. 5527890
GENERAL INFORMATION:
APPLICANT: Brown, Michael S.
APPLICANT: Briggs, Michael R.
APPLICANT: Wang, Xiaodong
APPLICANT: Goldstein, Joseph L.
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING
TO STEROL REGULATORY ELEMENT BINDING
TITLE OF INVENTION: PROTEINS
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: U.S.A.
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/131,365B
; FILING DATE: 01-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTSD:372/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-131-365B-48

Query Match 66.7%; Score 4; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05; Indels 0;
Matches 4; Conservative 0; Mismatches 0; Gaps 0;

Qy 2 HHHH 5
Db 1 HHHH 4

RESULT 32
US-08-469-486-47
; Sequence 47, Application US/08469486
; Patent No. 5739281
; GENERAL INFORMATION:
; APPLICANT: Thoesersen, Hans Christian
; APPLICANT: Holcset, Thor Las
; APPLICANT: Etzerodt, Michael
; TITLE OF INVENTION: Improved method for the refolding of
; TITLE OF INVENTION: proteins
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,486
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/192,060
; FILING DATE: February 4, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul T. Clark
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06363/002001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 542 5070
; TELEFAX: 617 542 8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear

; MOLECULE TYPE: peptide
US-08-469-486-47

Query Match 66.7%; Score 4; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05; Indels 0;
Matches 4; Conservative 0; Mismatches 0; Gaps 0;

Qy 2 HHHH 5
Db 1 HHHH 4

RESULT 33
US-08-460-343B-51
; Sequence 51, Application US/08460343B
; Patent No. 5741664
; GENERAL INFORMATION:
; APPLICANT: Marcus D. Ballinger and James A. Wells
; TITLE OF INVENTION: SUBSTITUTED VARIANTS CAPABLE OF CLEAVING
; TITLE OF INVENTION: SUBSTRATES CONTAINING DIBASIC RESIDUES
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,343B
; FILING DATE: 01-Jun-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/398028
; FILING DATE: 03-mar-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Kubinec, Jeffrey S.
; REGISTRATION NUMBER: 36,575
; REFERENCE/DOCKET NUMBER: P0936C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8228
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-460-343B-51

Query Match 66.7%; Score 4; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05; Indels 0;
Matches 4; Conservative 0; Mismatches 0; Gaps 0;

Qy 2 HHHH 5
Db 1 HHHH 4

RESULT 34
US-08-299-567-1
; Sequence 1, Application US/08299567
; Patent No. 5747033
; GENERAL INFORMATION:
; APPLICANT: Davis, et al.
; TITLE OF INVENTION: METHOD OF ENHANCING THE BIOLOGICAL
; TITLE OF INVENTION: ACTIVITY OF EPH FAMILY LIGANDS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Regeneron Pharmaceuticals, Inc.
STREET: 777 Old Saw Mill River Road
CITY: Tarrytown
STATE: New York
COUNTRY: U.S.A.
ZIP: 10591-6707
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/299,567
FILING DATE: 01-SEP-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kempler, Gail M.
REGISTRATION NUMBER: 32,143
REFERENCE/DOCKET NUMBER: REG 290
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-299-567-1

Query Match 66.7%; Score 4; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 1 HHHH 4

RESULT 35
US-08-398-028B-51
Sequence 51, Application US/08398028B
Patent No. 5780285
GENERAL INFORMATION:
APPLICANT: Marcus D. Ballinger and James A. Wells
TITLE OF INVENTION: SUBSTITUTED VARIANTS CAPABLE OF CLEAVING
NUMBER OF SEQUENCES: 74
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/398,028B
FILING DATE: 03-Mar-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kubinec, Jeffrey S.
REGISTRATION NUMBER: 36,575
REFERENCE/DOCKET NUMBER: P0936
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-8228
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 51:

SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-398-028B-51

Query Match 66.7%; Score 4; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 1 HHHH 4

RESULT 36
US-08-660-626-4
Sequence 4, Application US/08660626
Patent No. 5789655
GENERAL INFORMATION:
APPLICANT: Stanley B. Prusiner
APPLICANT: Glenn C. Telling
APPLICANT: Fred E. Cohen
APPLICANT: Michael R. Scott
TITLE OF INVENTION: TRANSGENIC ANIMALS EXPRESSING
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 2200 Sand Hill Road, Suite 100
CITY: Menlo Park
STATE: California
COUNTRY: USA
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/660,626
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Valeta Gregg
REGISTRATION NUMBER: 35,127
REFERENCE/DOCKET NUMBER: 07533/003001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 322-5070
TELEFAX: (415) 854-0875
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-660-626-4

Query Match 66.7%; Score 4; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 1 HHHH 4

RESULT 37
US-08-434-705B-17
Sequence 17, Application US/08434705B
Patent No. 5798258
GENERAL INFORMATION:
APPLICANT: Douglass Ph.D., James

TITLE OF INVENTION: CART PROTEIN AND DNA ENCODING THEREFOR
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis K. Shelton
STREET: 1420 Fifth Avenue, Suite 2800
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/434,705B
FILING DATE: May 4, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Shelton, Dennis K.
REGISTRATION NUMBER: 26,997
REFERENCE/DOCKET NUMBER: OHSU18446
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 682-8100
TELEFAX: (206) 224-0779
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
JS-08-434-705B-17

Query Match 66.7%; Score 4; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2y 2 HHHH 5
b 1 HHHH 4

RESULT 38
S-08-470-837-22
Sequence 22, Application US/08470837
Patent No. 5800811
GENERAL INFORMATION:
APPLICANT: Nimmi, Marcel E.
APPLICANT: Hall, Frederick L.
APPLICANT: Tuan, Tai-Lan
APPLICANT: Wu, Lingtao
APPLICANT: Cheung, David T.
TITLE OF INVENTION: Transforming Growth Factor B Fusion
TITLE OF INVENTION: and
TITLE OF INVENTION: Their Use in Wound Healing
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant & Gould
STREET: 11150 Santa Monica Boulevard, Suite 400
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90025-3395
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,837
FILING DATE:

CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharp, Janice A.
REGISTRATION NUMBER: 34,051
REFERENCE/DOCKET NUMBER: 30630-1US01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 310-445-1140
TELEFAX: 310-445-9031
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-470-837-22

Query Match 66.7%; Score 4; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
Db 1 HHHH 4

RESULT 39
US-08-472-244-14
Sequence 14, Application US/08472244
Patent No. 5821088
GENERAL INFORMATION:
APPLICANT: Darzins, Aldis
APPLICANT: Whitehead, Stephen
APPLICANT: Hruby, Dennis E.
TITLE OF INVENTION: Use of Gram-Positive Bacteria to Express
TITLE OF INVENTION: Recombinant Proteins
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, LLP
STREET: 699 Prince Street
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22314-3187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,244
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 016921-065
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-472-244-14

Query Match 66.7%; Score 4; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5

Db 1 HHHH 4

RESULT 40
US-08-504-265B-51
; Sequence 51, Application US/08504265B
; Patent No. 5837516
; GENERAL INFORMATION:
; APPLICANT: Marcus D. Ballinger and James A. Wells
; TITLE OF INVENTION: SUBSTITUTED VARIANTS CAPABLE OF CLEAVING
; TITLE OF INVENTION: SUBSTRATES CONTAINING BASIC RESIDUES
; NUMBER OF SEQUENCES: 90
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Winpatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/504,265B
; FILING DATE: 19-Jul-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/398028
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Kubinec, Jeffrey S.
; REGISTRATION NUMBER: 36,575
; REFERENCE/DOCKET NUMBER: P0936P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8228
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-504-265B-51

Query Match 66.7%; Score 4; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
Db 1 HHHH 4

RESULT 41
US-08-623-833B-16
; Sequence 16, Application US/08623833B
; Patent No. 5866683
; GENERAL INFORMATION:
; APPLICANT: SHIMURA, Kiyohito
; APPLICANT: KASAI, Kenichi
; APPLICANT: MATSUMOTO, Hiroyuki
; APPLICANT: TAKAMOTO, Hiroyoshi
; TITLE OF INVENTION: ISOELECTRIC POINT MARKERS FOR
; TITLE OF INVENTION: ISOELECTRIC FOCUSING WITH FLUORESCENCE DETECTION
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA

ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/623,833B
; FILING DATE: 29-MAR-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 0768663/1995
; FILING DATE: 31-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 271196/1995
; FILING DATE: 19-OCT-1995
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acid
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-623-833B-16

Query Match 66.7%; Score 4; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
Db 1 HHHH 4

RESULT 42
US-08-623-833B-56
; Sequence 56, Application US/08623833B
; Patent No. 5866683
; GENERAL INFORMATION:
; APPLICANT: SHIMURA, Kiyohito
; APPLICANT: KASAI, Kenichi
; APPLICANT: MATSUMOTO, Hiroyuki
; APPLICANT: TAKAMOTO, Hiroyoshi
; TITLE OF INVENTION: ISOELECTRIC POINT MARKERS FOR
; TITLE OF INVENTION: ISOELECTRIC FOCUSING WITH FLUORESCENCE DETECTION
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/623,833B
; FILING DATE: 29-MAR-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 0768663/1995
; FILING DATE: 31-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 271196/1995
; FILING DATE: 19-OCT-1995
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acid
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

JS-08-623-833B-56

Query Match 66.7%; Score 4; DB 2; Length 6;
Best Local Similarity 100.0%; Pred.No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 HHHH 5
b 2 HHHH 5

RESULT 43

JS-08-652-507-8
Sequence 8, Application US/08652507
Patent No. 5876691
GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: ANTIBODY AGAINST CARCINOEMBRYONIC ANTIGEN

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Nixon & Vanderhye, P.C.

STREET: 1100 No. 5876691th Glebe Road, 8th Floor

CITY: Arlington

STATE: Virginia

COUNTRY: USA

ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/652,507

FILING DATE: 02-Jul-1996

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Arthur R. Crawford

REGISTRATION NUMBER: 25,327

REFERENCE/DOCKET NUMBER: 117-211

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-816-4000

TELEFAX: 703-816-4100

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 6 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

S-08-652-507-8

Query Match 66.7%; Score 4; DB 2; Length 6;
Best Local Similarity 100.0%; Pred.No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 HHHH 5
b 1 HHHH 4

RESULT 44

S-08-668-123-48
Sequence 48, Application US/08668123
Patent No. 5891631
GENERAL INFORMATION:

APPLICANT: Brown, Michael S.

APPLICANT: Briggs, Michael R.

APPLICANT: Wang, Xiaodong

APPLICANT: Goldstein, Joseph L.

TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING

TITLE OF INVENTION: TO STEROL REGULATORY ELEMENT BINDING

NUMBER OF SEQUENCES: 64

CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: U.S.A.
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/668,123
FILING DATE: 14-JUN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/131,365
FILING DATE: 01-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UTSD:372/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-668-123-48

Query Match 66.7%; Score 4; DB 2; Length 6;
Best Local Similarity 100.0%; Pred.No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
Db 1 HHHH 4

RESULT 45

US-08-811-028-49
Sequence 49, Application US/08811028C
Patent No. 5891671
GENERAL INFORMATION:

APPLICANT: SUZUKI, Yuji

APPLICANT: MAGOTA, Koji

APPLICANT: MASUDA, Toyofumi

TITLE OF INVENTION: METHOD FOR CLEAVING CHIMERIC ENZYME USING PROCESSING

TITLE OF INVENTION: ENZYME

FILE REFERENCE: 001560-294

CURRENT APPLICATION NUMBER: US/08/811,028C

CURRENT FILING DATE: 1987-03-04

EARLIER FILING DATE: 1996-03-04

NUMBER OF SEQ ID NOS: 54

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 49

LENGTH: 6

TYPE: PRI

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence:exemplary

OTHER INFORMATION: linker peptide

US-08-811-028-49

Query Match 66.7%; Score 4; DB 2; Length 6;
Best Local Similarity 100.0%; Pred.No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
||||
DB 1 HHHH 4

RESULT 46
US-08-469-658-47
; Sequence 47, Application US/08469658
; Patent No. 5917018
; GENERAL INFORMATION:
; APPLICANT: Th egersen, Hans Christian
; APPLICANT: Holtet, Thor Las
; APPLICANT: Etzerodt, Michael
; TITLE OF INVENTION: IMPROVED METHOD FOR THE REFOOLDING OF
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,658
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/192,060
; FILING DATE: February 4, 1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul T. Clark
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06363/002002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 542 5070
; TELEFAX: 617 542 8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: linear
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-469-658-47

Query Match 66.7%; Score 4; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
||||
DB 1 HHHH 4

RESULT 47
US-09-086-201-17
; Sequence 17, Application US/09086201
; Patent No. 5969103
; GENERAL INFORMATION:
; APPLICANT: Douglass Ph.D., James
; TITLE OF INVENTION: CART PROTEIN AND DNA ENCODING THEREFOR
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dennis K. Shelton
; STREET: 1420 Fifth Avenue, Suite 2800
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/086,201
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/434,705
; FILING DATE: May 4, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Shelton, Dennis K.
; REGISTRATION NUMBER: 26,997
; REFERENCE/DOCKET NUMBER: OHSU18446
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 682-8100
; TELEFAX: (206) 224-0779
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
US-09-086-201-17

Query Match 66.7%; Score 4; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
||||
DB 1 HHHH 4

RESULT 48
US-08-591-196-57
; Sequence 57, Application US/08591196
; Patent No. 597316
; GENERAL INFORMATION:
; APPLICANT: Chatterjee, Malaya
; APPLICANT: Foon, Kenneth A.
; APPLICANT: Chatterjee, Sunil K.
; TITLE OF INVENTION: MONOCLONAL ANTIBODY 1A7 AND USE FOR THE
; TITLE OF INVENTION: TREATMENT OF MELANOMA AND SMALL CELL CARCINOMA
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,196
; FILING DATE: 16-JAN-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Schiff, J. Michael

REGISTRATION NUMBER: 40,253
 REFERENCE/DOCKET NUMBER: 30414-20002.20
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 813-5600
 TELEFAX: (415) 494-0792
 TELEX: 706141
 INFORMATION FOR SEQ ID NO: 57:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 6 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 JS-08-591-196-57

Query Match 66.7%; Score 4; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 HHHH 5
 ||||
 b 1 HHHH 4

RESULT 49
 JS-08-996-139-18
 Sequence 18, Application US/08996139
 Patent No. 6017729
 GENERAL INFORMATION:
 APPLICANT: Anderson, Dirk M.
 APPLICANT: Galibert, Laurent
 APPLICANT: Maraskovsky, Eugene
 TITLE OF INVENTION: Receptor Activator of NF-kappaB
 NUMBER OF SEQUENCES: 19
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Immunex Corporation, Law Department
 STREET: 51 University Street
 CITY: Seattle
 STATE: WA
 COUNTRY: USA
 ZIP: 98101

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: Apple Power Macintosh
 OPERATING SYSTEM: Apple Operating System 7.5.5
 SOFTWARE: Microsoft Word for Power Macintosh 6.0.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/996,139
 FILING DATE: 22 DECEMBER 1997

CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: USSN 60/064,671
 FILING DATE: 14 OCTOBER 1997

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: USSN 08/813,509
 FILING DATE: 07 MARCH 1997

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: USSN 08/772,330
 FILING DATE: 23 DECEMBER 1996

ATTORNEY/AGENT INFORMATION:
 NAME: Perkins, Patricia Anne
 REGISTRATION NUMBER: 34,693
 REFERENCE/DOCKET NUMBER: 2851-A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206)587-0430
 TELEFAX: (206)233-0644

INFORMATION FOR SEQ ID NO: 18:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 6 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 JS-08-996-139-18

Query Match 66.7%; Score 4; DB 3; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
 ||||
 Db 1 HHHH 4

RESULT 50
 US-08-481-435-42
 ; Sequence 42, Application US/08481435
 ; Patent No. 6027906
 ; GENERAL INFORMATION:
 ; APPLICANT: Balganes, Tanjore S
 ; APPLICANT: Town, Christine
 ; TITLE OF INVENTION: No. 6027906el Polypeptides
 ; NUMBER OF SEQUENCES: 42
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: White & Case
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: United States
 ; ZIP: 10036-2787
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/481,435
 ; FILING DATE: 10-JUL-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: IN 580/MAS/94
 ; FILING DATE: 01-JUL-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: SE 9404072-2
 ; FILING DATE: 24-NOV-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sterner, Richard J.
 ; REGISTRATION NUMBER: 35,372
 ; REFERENCE/DOCKET NUMBER: 1103326-151
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 819-8783
 ; TELEFAX: (212) 354-8113
 ; INFORMATION FOR SEQ ID NO: 42:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 6 amino acids
 ; TYPE: amino acids
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-481-435-42

Query Match 66.7%; Score 4; DB 3; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
 ||||
 Db 1 HHHH 4

RESULT 51
 US-08-772-440-28
 ; Sequence 28, Application US/08772440
 ; Patent No. 6046158
 ; GENERAL INFORMATION:
 ; APPLICANT: Arizumi, Kiyoshi
 ; APPLICANT: Takashima, Akira
 ; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE

;; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES

;; TITLE OF INVENTION: THEREOF

;; NUMBER OF SEQUENCES: 42

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: Arnold, White & Durkee

;; STREET: P.O. Box 4433

;; CITY: Houston

;; STATE: Texas

;; COUNTRY: USA

;; ZIP: 77210

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: PatentIn Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/772,440

;; FILING DATE: CONCURRENTLY HERewith

;; CLASSIFICATION: 435

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Parker, David L.

;; REGISTRATION NUMBER: 32,165

;; REFERENCE/DOCKET NUMBER: UTXD:493

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: 512/418-3000

;; TELEFAX: 512/474-7577

;; INFORMATION FOR SEQ ID NO: 28:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 6 amino acids

;; TYPE: amino acid

;; STRANDEDNESS:

;; TOPOLOGY: linear

;; US-08-772-440-28

Query Match 66.7%; Score 4; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 1 HHHH 4

RESULT 52
US-09-267-031-18
; Sequence 18, Application US/09267031
; Patent No. 6137031
; GENERAL INFORMATION:
; APPLICANT: Zhang, Yuelin
; APPLICANT: Kinkema, Mark
; APPLICANT: Dong, Xinlian
; APPLICANT: Rorald, Pamela
; APPLICANT: Chern, MawSheng
; APPLICANT: The Regents of the University of California
; APPLICANT: Duke University
; TITLE OF INVENTION: DNA Binding Proteins That Interact With NPR1
; FILE REFERENCE: 023070-092500US
; CURRENT APPLICATION NUMBER: US/09/267,031
; CURRENT FILING DATE: 1999-03-11
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: (His)6-tag
US-09-267-031-18

Query Match 66.7%; Score 4; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 1 HHHH 4

RESULT 53
US-09-398-341-1
; Sequence 1, Application US/09398341
; Patent No. 6146842
; GENERAL INFORMATION:
; APPLICANT: Josiah, Serene
; APPLICANT: Boisclair, Michael
; TITLE OF INVENTION: High-Throughput Screening Assays Utilizing Metal-
; TITLE OF INVENTION: Chelate Capture
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: ONE POST OFFICE SQUARE
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPad
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/398,341
; FILING DATE: 17-SEP-1999
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-116.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-398-341-1

Query Match 66.7%; Score 4; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 1 HHHH 4

RESULT 54
US-09-031-168-4
; Sequence 4, Application US/09031168
; Patent No. 6150583
; GENERAL INFORMATION:
; APPLICANT: Stanley B. Prusiner
; APPLICANT: Glenn C. Telling
; APPLICANT: Fred E. Cohen
; APPLICANT: Michael R. Scott
; TITLE OF INVENTION: TRANSGENIC ANIMALS EXPRESSING
; TITLE OF INVENTION: EPIOTOPE-TAGGED PROTEINS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: California
; COUNTRY: USA


```
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/031.168
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/660,626
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Valeta Gregg
REGISTRATION NUMBER: 35,127
REFERENCE/DOCKET NUMBER: 07532/003001
TELEPHONE: (415) 322-5070
TELEFAX: (415) 854-0875
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-031-168-4

Query Match 66.7%; Score 4; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 1 HHHH 4

RESULT 55
US-08-789-333F-99
Sequence 99, Application US/08789333F
Patent No. 6153380
GENERAL INFORMATION:
APPLICANT: No. 6153380an, Garry P
APPLICANT: Rothenberg, S. M.
TITLE OF INVENTION: METHODS FOR SCREENING FOR TRANSDOMINANT INTRACELLULAR
TITLE OF INVENTION: EFFECTOR PEPTIDES AND RNA MOLECULES
FILE REFERENCE: A642601DJBRMSDS
CURRENT APPLICATION NUMBER: US/08/789,333F
CURRENT FILING DATE: 1997-01-23
PRIOR APPLICATION NUMBER: 08/589,108
PRIOR FILING DATE: 1996-01-23
PRIOR APPLICATION NUMBER: 08/589,911
PRIOR FILING DATE: 1996-01-23
NUMBER OF SEQ ID NOS: 102
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 99
LENGTH: 6
TYPE: PPT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: histidine tag
US-08-789-333F-99

Query Match 66.7%; Score 4; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 1 HHHH 4

RESULT 56
US-08-995-659-18
Sequence 18, Application US/08995659
Patent No. 6242213
GENERAL INFORMATION:
APPLICANT: Anderson, Dirk M.
APPLICANT: Galibert, Laurent
APPLICANT: Maraskovsky, Eugene
TITLE OF INVENTION: Ligand for Receptor Activator of NF-kappaB
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation, Law Department
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Power Macintosh 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/995,659
FILING DATE: 22 DECEMBER 1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 60/064,671
FILING DATE: 14 OCTOBER 1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/813,509
FILING DATE: 07 MARCH 1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/772,330
FILING DATE: 23 DECEMBER 1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2852-A
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-995-659-18

Query Match 66.7%; Score 4; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 1 HHHH 4

RESULT 57
US-09-142-334-26
Sequence 26, Application US/09142334
Patent No. 6268485
GENERAL INFORMATION:
APPLICANT: Farries, Timothy C.
APPLICANT: Harrison, Richard A.
TITLE OF INVENTION: Down-Regulation Resistant C3 Convertase
FILE REFERENCE: 4-30443/A/IMU/PCT
CURRENT APPLICATION NUMBER: US/09/142,334
CURRENT FILING DATE: 1999-04-15
```

EARLIER APPLICATION NUMBER: PCT/GB97/00603
EARLIER FILING DATE: 1997-03-04
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 26
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: amino acid
OTHER INFORMATION: Insertion
US-09-142-334-26

Query Match 66.7%; Score 4; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
Db 1 HHHH 4

RESULT 58

US-09-215-649A-18
Sequence 18, Application US/09215649A
Patent No. 6271349
GENERAL INFORMATION:

APPLICANT: Anderson, Dirk M.
Galibert, Laurent
Marakovsky, Eugene
TITLE OF INVENTION: Receptor Activator of NF-kappaB
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation, Law Department
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Power Macintosh 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/215.649A
FILING DATE: 17-Dec-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/996,139
FILING DATE: <Unknown>
APPLICATION NUMBER: USSN 08/813,509
FILING DATE: 07 MARCH 1997
APPLICATION NUMBER: USSN 08/772,330
FILING DATE: 23 DECEMBER 1996
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2851-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 18:

US-09-215-649A-18

Query Match 66.7%; Score 4; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 HHHH 5
Db 1 HHHH 4

RESULT 59

US-09-140-084-7
Sequence 7, Application US/09140084A
Patent No. 6300065
GENERAL INFORMATION:

APPLICANT: Kieke, et al.
TITLE OF INVENTION: Yeast Cell Surface Display of Proteins and Uses Thereof
FILE REFERENCE: D6061CIP2
CURRENT APPLICATION NUMBER: US/09/140.084A
CURRENT FILING DATE: 1998-08-26
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 6
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Epitope Tag
US-09-140-084-7

Query Match 66.7%; Score 4; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
Db 1 HHHH 4

RESULT 60

US-09-382-950-4
Sequence 4, Application US/09382950
Patent No. 6303337
GENERAL INFORMATION:

APPLICANT: Rothschild, Kenneth
APPLICANT: Gite, Sadanand
APPLICANT: Olejnik, Jerzy
TITLE OF INVENTION: N-Terminal and C-Terminal Markers in Nascent Proteins
FILE REFERENCE: AMBER-03879
CURRENT APPLICATION NUMBER: US/09/382.950
CURRENT FILING DATE: 1999-08-25
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial/Unknown
FEATURE:
NAME/KEY: misc feature
LOCATION: ().()
OTHER INFORMATION: Synthetic
US-09-382-950-4

Query Match 66.7%; Score 4; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
Db 1 HHHH 4

RESULT 61

US-09-382-736B-5
Sequence 5, Application US/09382736B
Patent No. 6306628

GENERAL INFORMATION:

APPLICANT: Rothechild, Kenneth
APPLICANT: Gite, Sadanand
APPLICANT: Olejnik, Jerzy

TITLE OF INVENTION: Methods for the Detection, Analysis and Isolation of Nascent Protein

FILE REFERENCE: AMBER-03951

CURRENT APPLICATION NUMBER: US/09/382.736B

CURRENT FILING DATE: 1999-08-25

NUMBER OF SEQ ID NOS: 18

SOFTWARE: PatentIn version 3.0

SEQ ID NO 5

LENGTH: 6

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Synthetic

JS-09-382-736B-5

Query Match

Best Local Similarity 66.7%; Score 4; DB 4; Length 6;

Mismatches 0; Indels 0; Gaps 0;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

2 HHHH 5

||||

1 HHHH 4

Db

RESULT 62

JS-08-977-378-2

Sequence 2, Application US/08977378

Patent No. 6309842

GENERAL INFORMATION:

APPLICANT: Dower, William J.

Gates, Christopher M.

Heinkel, Gregory L.

Lalonde, Guy

Mattheakis, Larry C.

Paddon, Christopher J.

Schaltz, Peter J.

TITLE OF INVENTION: Use of Modified Tethers in Screening Compound Libraries

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/977,378

FILING DATE: 24-NOV-6309842-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/758,307

FILING DATE: 03-DEC-1996

ATTORNEY/AGENT INFORMATION:

NAME: Liebeschuetz, Joe

REGISTRATION NUMBER: 37,505

REFERENCE/DOCKET NUMBER: 16528A-018010US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 6 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-977-378-2

Query Match 66.7%; Score 4; DB 4; Length 6;

Best Local Similarity 100.0%; Pred. No. 2.5e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5

||||

1 HHHH 4

Db

RESULT 63

US-09-353-555-2

Sequence 2, Application US/09353555

Patent No. 6329209

GENERAL INFORMATION:

APPLICANT: Wagner, Peter

APPLICANT: No. 6329209k, Steffen

APPLICANT: Ault-Riche, Dana

APPLICANT: Itin, Christian

TITLE OF INVENTION: Arrays of Protein-Capture Agents and Methods of Use

TITLE OF INVENTION: Thereof

FILE REFERENCE: 24406-0006

CURRENT APPLICATION NUMBER: US/09/353,555

CURRENT FILING DATE: 1999-07-14

EARLIER APPLICATION NUMBER: US 09/115,455

EARLIER FILING DATE: 1998-07-14

NUMBER OF SEQ ID NOS: 2

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2

LENGTH: 6

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: histidine tag

US-09-353-555-2

Query Match

Best Local Similarity 66.7%; Score 4; DB 4; Length 6;

Mismatches 0; Indels 0; Gaps 0;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5

||||

1 HHHH 4

Db

RESULT 64

US-08-868-452-22

Sequence 22, Application US/08868452C

Patent No. 6352972

GENERAL INFORMATION:

APPLICANT: Marcel E. Nimni

APPLICANT: Frederick L. Hall

APPLICANT: Lingtao Wu

APPLICANT: Bo Han

APPLICANT: Edwin Shors

TITLE OF INVENTION: BONE MORPHOGENETIC PROTEINS AND THEIR

TITLE OF INVENTION: USE IN BONE GROWTH

FILE REFERENCE: 17972-11

CURRENT APPLICATION NUMBER: US/08/868,452C

CURRENT FILING DATE: 1997-06-03

NUMBER OF SEQ ID NOS: 51

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 22

LENGTH: 6

TYPE: PRT

ORGANISM: Human

US-08-868-452-22

Query Match

Best Local Similarity 66.7%; Score 4; DB 4; Length 6;

Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
Db 1 HHHH 4

RESULT 65

US-09-000-094-47
; Sequence 47, Application US/09000094
; Patent No. 6365160
; GENERAL INFORMATION:
; APPLICANT: WEBB, Elizabeth Ann
; MARGGETTS, Mary Brigid
; COX, John Cooper
; FRAZER, Ian
; MCMILLAN, Nigel Alan John
; WILLIAMS, Mark Philip
; MOLONEY, Margaret Bridget
; Holland
; EDWARDS, Stirling John

TITLE OF INVENTION: PAPILLOMAVIRUS POLYPROTEIN CONSTRUCTS
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY & LARDNER
STREET: 3000 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/000.094
FILING DATE: 21-Apr-1998
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/AU96/00473
FILING DATE: 26-JUL-1996
APPLICATION NUMBER: AU PN 4439/95
FILING DATE: 27-JUL-1995

ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 017227/0137
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399

INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 47:

US-09-000-094-47
Query Match 66.7%; Score 4; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
Db 1 HHHH 4

RESULT 66

US-09-058-483-7

; Sequence 7, Application US/09058483A
; Patent No. 6365347
; GENERAL INFORMATION:
; APPLICANT: Murray, Andrew W.
; APPLICANT: Smith, Dana L.
; APPLICANT: Sorger, Peter K.
; APPLICANT: No. 6365347man, Thea C.

TITLE OF INVENTION: METHODS FOR IDENTIFYING DISRUPTORS OF BIOLOGICAL
FILE REFERENCE: 30432.1US11
CURRENT APPLICATION NUMBER: US/09/058.483A
CURRENT FILING DATE: 1998-04-10
EARLIER APPLICATION NUMBER: 08/835,727
EARLIER FILING DATE: 1997-04-11
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 6
TYPE: PPT
ORGANISM: amino acid sequence of Staphylococcus aureus nuclease protei

US-09-058-483-7

Query Match 66.7%; Score 4; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
Db 1 HHHH 4

RESULT 67

US-09-140-201-8
; Sequence 8, Application US/09140201
; Patent No. 6372425
; GENERAL INFORMATION:
; APPLICANT: KELLER, P.

TITLE OF INVENTION: LARGE SCALE AFFINITY CHROMATOGRAPHY OF
TITLE OF INVENTION: MACROMOLECULES
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:

ADDRESSEE: JACK TRIBBLE
STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE
CITY: RAHWAY
STATE: NJ
COUNTRY: USA
ZIP: 07065

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/140.201
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/329,749
FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: TRIBBLE, JACK
REGISTRATION NUMBER: 32,633
REFERENCE/DOCKET NUMBER: 18780
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-5321
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide

S-09-140-201-8

Query Match 66.7%; Score 4; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 HHHH 5
|
|
|
|
b 1 HHHH 4

RESULT 68

US-09-039-780A-106
Sequence 106, Application US/09039780A
Patent No. 6376248

GENERAL INFORMATION:

APPLICANT: HAWLEY-NELSON, PAMELA

LAN, JIANQING

SHIH, POJEN

JESSE, JOEL A.

SCHIFFERLI, KEVIN P.

GEBVEYHU, GUILIAT

TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSFECTIONS

NUMBER OF SEQUENCES: 120

CORRESPONDENCE ADDRESS:

ADDRESSEE: GREENLEE, WINNER & SULLIVAN

STREET: 5370 MANHATTAN CIRCLE, SUITE 201

CITY: BOULDER

STATE: CO

COUNTRY: US

ZIP: 80303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/039,780A

FILING DATE: 16-Mar-1998

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: SULLIVAN, SALLY A.

REGISTRATION NUMBER: 32,064

REFERENCE/DOCKET NUMBER: 32-95C

TELECOMMUNICATION INFORMATION:

TELEPHONE: (303)499-8080

TELEFAX: (303)499-8089

INFORMATION FOR SEQ ID NO: 106:

SEQUENCE CHARACTERISTICS:

LENGTH: 6 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

SEQUENCE DESCRIPTION: SEQ ID NO: 106:

US-09-039-780A-106

Query Match 66.7%; Score 4; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 HHHH 5
|
|
|
|
b 1 HHHH 4

RESULT 69

US-09-291-170A-14

Sequence 14, Application US/09291170A

Patent No. 6410687

GENERAL INFORMATION:

APPLICANT: Vale, Ronald D.
APPLICANT: Hartman, James J.
APPLICANT: The Regents of the University of California
TITLE OF INVENTION: Assays for the Detection of Microtubule
TITLE OF INVENTION: Depolymerization Inhibitors
FILE REFERENCE: 18557B-0005100S
CURRENT APPLICATION NUMBER: US/09/291,170A
CURRENT FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US 60/081,734
PRIOR FILING DATE: 1998-04-14
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 14
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: (His)6 or 6xHis
tag
US-C9-291-170A-14

Query Match 66.7%; Score 4; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
|
|
|
|
Db 1 HHHH 4

RESULT 70

US-09-577-780-18

Sequence 18, Application US/G9577780

Patent No. 6419929

GENERAL INFORMATION:

APPLICANT: Anderson, Dirk M.

Galibert, Laurent

Maraskovsky, Eugene

TITLE OF INVENTION: Ligand for Receptor Activator of NF-kappaB

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: Immunex Corporation, Law Department

STREET: 51 University Street

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: Apple Power Macintosh

OPERATING SYSTEM: Apple Operating System 7.5.5

SOFTWARE: Microsoft Word for Power Macintosh 6.0.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/577,780

FILING DATE: 24-May-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/995,659

FILING DATE: <Unknown>

APPLICATION NUMBER: USSN 08/813,509

FILING DATE: 07 MARCH 1997

APPLICATION NUMBER: USSN 08/772,330

FILING DATE: 23 DECEMBER 1996

ATTORNEY/AGENT INFORMATION:

NAME: Perkins, Patricia Anne

REGISTRATION NUMBER: 34,693

REFERENCE/DOCKET NUMBER: 2852-A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206)587-0430

TELEFAX: (206)233-0644

INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:

LENGTH: 6 amino acids

```
;
;      TYPE: amino acid
;      TOPOLOGY: linear
;      MOLECULE TYPE: protein
;      SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-577-780-18
Query Match      66.7%; Score 4; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 HHHH 5
Db      1 HHHH 4

RESULT 71
US-09-724-297-7
; Sequence 7, Application US/09724297
; Patent No. 6423538
; GENERAL INFORMATION:
; APPLICANT: The Board of Trustees of the University of Illinois
; APPLICANT: Wittrup, et al.
; TITLE OF INVENTION: Yeast Cell Surface Display of Proteins and Uses Thereof
; FILE REFERENCE: 97-99C
; CURRENT APPLICATION NUMBER: US/09/724,297
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 09/009,388
; PRIOR FILING DATE: 1998-01-20
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 6
; TYPE: PRT
; ORGANISM: unknown
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (..)
; OTHER INFORMATION: Epitope tag
US-09-724-297-7
Query Match      66.7%; Score 4; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 HHHH 5
Db      1 HHHH 4

RESULT 72
US-09-367-309A-4
; Sequence 4, Application US/09367309A
; Patent No. 6428807
; GENERAL INFORMATION:
; APPLICANT: MACFARLAN, RODERICK I.
; APPLICANT: MALLIAROS, JIM
; TITLE OF INVENTION: CHELATING IMMUNOSTIMULATING COMPLEXES
; FILE REFERENCE: 017227/0149
; CURRENT APPLICATION NUMBER: US/09/367,309A
; CURRENT FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: PCT/AU98/00080
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: AU PO 5178
; PRIOR FILING DATE: 1997-02-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: 6-His tag
US-09-367-309A-4
Query Match      66.7%; Score 4; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 HHHH 5
Db      1 HHHH 4

RESULT 73
US-09-724-884-14
; Sequence 14, Application US/09724884
; Patent No. 6429304
; GENERAL INFORMATION:
; APPLICANT: Vale, Ronald D.
; APPLICANT: Hartman, James J.
; TITLE OF INVENTION: Assays for the Detection of Microtubule
; TITLE OF INVENTION: Depolymerization Inhibitors
; FILE REFERENCE: 18557B-000510US
; CURRENT APPLICATION NUMBER: US/09/724,884
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/291,170
; PRIOR FILING DATE: 1999-04-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: (His)6 or 6xHis
US-09-724-884-14
Query Match      66.7%; Score 4; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 HHHH 5
Db      1 HHHH 4

RESULT 74
US-09-529-279-6
; Sequence 6, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529,279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-529-279-6
```

Query Match 66.7%; Score 4; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 1 HHHH 4

RESULT 75
US-09-313-942-1
Sequence 1, Application US/09313942
Patent No. 6472179
GENERAL INFORMATION:
APPLICANT: REGENERON PHARMACEUTICALS, INC.
TITLE OF INVENTION: RECEPTOR BASED ANTAGONISTS, AND METHODS OF MAKING
TITLE OF INVENTION: AND USING
FILE REFERENCE: REG 203-A
CURRENT APPLICATION NUMBER: US/09/313,942
CURRENT FILING DATE: 1999-05-19
PRIOR APPLICATION NUMBER: 09/313,942
PRIOR FILING DATE: 1999-05-19
PRIOR APPLICATION NUMBER: 60/101,858
PRIOR FILING DATE: 1998-09-25
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 1
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Tag sequence
US-09-313-942-1

Query Match 66.7%; Score 4; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 1 HHHH 4

RESULT 76
US-09-353-215-6
Sequence 6, Application US/09353215
Patent No. 6475808
GENERAL INFORMATION:
APPLICANT: Wagner, Peter
APPLICANT: Ault-Riche, Dana
APPLICANT: No. 6475808k, Steffen
APPLICANT: Itin, Christian
TITLE OF INVENTION: Arrays of Proteins and Methods of Use Thereof
FILE REFERENCE: 24406-0004P1
CURRENT APPLICATION NUMBER: US/09/353,215
CURRENT FILING DATE: 1999-07-14
EARLIER APPLICATION NUMBER: US 09/115,455
EARLIER FILING DATE: 1998-07-14
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 6
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: histidine tag
US-09-353-215-6

Query Match 66.7%; Score 4; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5

DB 1 HHHH 4

RESULT 77
US-09-577-800-18
Sequence 18, Application US/09577800
Patent No. 6479635
GENERAL INFORMATION:
APPLICANT: Anderson, Dirk M.
APPLICANT: Galibert, Laurent
APPLICANT: Maraskovsky, Eugene
TITLE OF INVENTION: Receptor Activator of NF-kappaB
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation, Law Department
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Power Macintosh 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/577,800
FILING DATE: 24-MAY-2000
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/996,139
FILING DATE: 22 DECEMBER 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 60/064,671
FILING DATE: 14 OCTOBER 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/813,509
FILING DATE: 07 MARCH 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/772,330
FILING DATE: 23 DECEMBER 1996
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2851-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)233-0644
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-577-800-18

Query Match 66.7%; Score 4; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 1 HHHH 4

RESULT 78
US-09-667-422-7
Sequence 7, Application US/09667422
Patent No. 6482611
GENERAL INFORMATION:
APPLICANT: Cortright, Daniel
APPLICANT: Krause, James

; TITLE OF INVENTION: Human Capsaicin Receptor and Uses Thereof
; FILE REFERENCE: HCR
; CURRENT APPLICATION NUMBER: US/09/667,422
; CURRENT FILING DATE: 2001-06-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:His6x epitope
US-09-667-422-7

Query Match 66.7%; Score 4; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
|
|
|
|
Db 1 HHHH 4

RESULT 79
US-09-315-926A-81
; Sequence 81, Application US/09315926A
; Patent No. 6498027
; GENERAL INFORMATION:
; APPLICANT: Es van, Helmut
; APPLICANT: Havenga, Menzo
; APPLICANT: Verlinden, Stefan
; TITLE OF INVENTION: TARGETED DELIVERY THROUGH A CATIONIC AMINO ACID TRANSPORTER
; FILE REFERENCE: 2183-4080US
; CURRENT APPLICATION NUMBER: US/09/315,926A
; CURRENT FILING DATE: 1999-05-20
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: EP 98201593.3
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: EP 98201693.3
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 81
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Description of Artificial Sequence: phage
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(6)
; OTHER INFORMATION: /note="hCAT1 amino acid sequence"
US-09-315-926A-81

Query Match 66.7%; Score 4; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
|
|
|
|
Db 1 HHHH 4

RESULT 80
US-09-731-558-11
; Sequence 11, Application US/09731558
; Patent No. 6503717
; GENERAL INFORMATION:
; APPLICANT: Case, Casey Christopher
; APPLICANT: Liu, Qiang
; APPLICANT: Rebar, Edward J.
; APPLICANT: Sangamo Biosciences, Inc.
; TITLE OF INVENTION: Methods of Using Randomized Libraries of Zinc Finger
; TITLE OF INVENTION: Proteins for the Identification of Gene Function

; FILE REFERENCE: 019496-003210US
; CURRENT APPLICATION NUMBER: US/09/731,558
; CURRENT FILING DATE: 2000-12-06
; PRIOR APPLICATION NUMBER: US 09/456,100
; PRIOR FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:6xHis tag
US-09-731-558-11

Query Match 66.7%; Score 4; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
|
|
|
|
Db 1 HHHH 4

RESULT 81
US-09-280-030-61
; Sequence 61, Application US/09280030A
; Patent No. 6506595
; GENERAL INFORMATION:
; APPLICANT: Sato, Seiji
; APPLICANT: Higashikuni, Naohiko
; APPLICANT: Kudo, Toshiyuki
; APPLICANT: Kondo, Masaaki
; TITLE OF INVENTION: DNAs ENCODING NEW FUSION PROTEINS AND PROCESSES FOR THE
; TITLE OF INVENTION: PREPARING USEFUL POLYPEPTIDES THROUGH EXPRESSION OF THE
; FILE REFERENCE: 382.1026
; CURRENT APPLICATION NUMBER: US/09/280,030A
; CURRENT FILING DATE: 1999-03-26
; EARLIER APPLICATION NUMBER: JP10-87339/1998
; EARLIER FILING DATE: 1998-03-31
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Designated is
; OTHER INFORMATION: a tag for separation/purification of a fusion
; OTHER INFORMATION: protein
US-09-280-030-61

Query Match 66.7%; Score 4; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
|
|
|
|
Db 1 HHHH 4

RESULT 82
US-09-466-496-18
; Sequence 18, Application US/09466496
; Patent No. 6528482
; GENERAL INFORMATION:
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Galibert, Laurent
; APPLICANT: Maraskovsky, Eugene
; TITLE OF INVENTION: Receptor Activator of NF-kappaB
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Immunex Corporation, Law Department
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Power Macintosh 6.0.1
CURRENT APPLICATION DATA: US/09/466,496
FILING DATE: 17-Dec-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/996,139
FILING DATE: 22 DECEMBER 1997
APPLICATION NUMBER: USN 60/064,671
FILING DATE: 14 OCTOBER 1997
APPLICATION NUMBER: USN 08/813,509
FILING DATE: 07 MARCH 1997
APPLICATION NUMBER: USN 08/772,330
FILING DATE: 23 DECEMBER 1996
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2851-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)233-0644
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-466-496-18
Query Match 66.7%; Score 4; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;
Qy 2 HHHH 5
Db 1 HHHH 4
RESULT 83
US-09-648-569A-39
Sequence 39, Application US/09648569A
Patent No. 6531122
GENERAL INFORMATION:
APPLICANT: Pedersen, A.H., et al.
TITLE OF INVENTION: Interferon-Beta Variants and Conjugates
FILE REFERENCE: 0202u8810
CURRENT APPLICATION NUMBER: US/09/648,569A
CURRENT FILING DATE: 2000-08-25
NUMBER OF SEQ ID NOS: 45
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 39
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: peptide tag
US-09-648-569A-39
Query Match 66.7%; Score 4; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;

Qy 2 HHHH 5
Db 1 HHHH 4
RESULT 84
US-09-871-856-18
Sequence 18, Application US/09871856
Patent No. 6537763
GENERAL INFORMATION:
APPLICANT: Anderson, Dirk M.
Galibert, Laurent
Maraskovsky, Eugene
TITLE OF INVENTION: Receptor Activator of NF-kappaB
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation, Law Department
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Power Macintosh 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/871,856
FILING DATE: 31-May-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/996,139
FILING DATE: <Unknown>
APPLICATION NUMBER: USN 08/813,509
FILING DATE: 07 MARCH 1997
APPLICATION NUMBER: USN 08/772,330
FILING DATE: 23 DECEMBER 1996
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2851-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)233-0644
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-871-856-18
Query Match 66.7%; Score 4; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;
Qy 2 HHHH 5
Db 1 HHHH 4
RESULT 85
US-10-158-895-6
Sequence 6, Application US/10158895
Patent No. 6551840
GENERAL INFORMATION:
APPLICANT: ONO, KOICHIRO
APPLICANT: OHTOMO, TOSHIHIKO
APPLICANT: TSUCHIYA, MASAYUKI
TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES

; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-158-895-6

Query Match 66.7%; Score 4; DB 4; Length 6;
Best Local Similarity 100.0%; Pred.No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
|||||
Db 1 HHHH 4

RESULT 86
US-09-904-196B-9
; Sequence 9, Application US/09904196B
; Patent No. 6555660
; GENERAL INFORMATION:
; APPLICANT: NISSEN, TORBEN LAUESGAARD
; APPLICANT: ANDERSEN, KIM VILBOUR
; APPLICANT: HANSEN, CHRISTIAN KARSTEN
; APPLICANT: MIKKELSEN, JAN MOLLER
; TITLE OF INVENTION: G-CSF CONJUGATES
; FILE REFERENCE: 31-00070005
; CURRENT APPLICATION NUMBER: US/09/904,196B
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US/09/760,008
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/176,376
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 60/189,506
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 60/215,644
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DK PA 2000 00024
; PRIOR FILING DATE: 2000-01-10
; PRIOR APPLICATION NUMBER: DK PA 2000 00341
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: DK PA 2000 00943
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-904-196B-9

Query Match 66.7%; Score 4; DB 4; Length 6;
Best Local Similarity 100.0%; Pred.No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
|||||

Db 1 HHHH 4

RESULT 87
US-09-871-291-18
; Sequence 18, Application US/09871291
; Patent No. 6562948
; GENERAL INFORMATION:
; APPLICANT: Anderson, Dirk M.
; Galibert, Laurent
; Maraskovsky, Eugene
; TITLE OF INVENTION: Receptor Activator of NF-kappaB
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation, Law Department
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for Power Macintosh 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/871,291
; FILING DATE: 30-May-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/996,139
; FILING DATE: <Unknown>
; APPLICATION NUMBER: USSN 08/813,509
; FILING DATE: 07 MARCH 1997
; APPLICATION NUMBER: USSN 08/772,330
; FILING DATE: 23 DECEMBER 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,693
; REFERENCE/DOCKET NUMBER: 2851-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)233-0644
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-871-291-18

Query Match 66.7%; Score 4; DB 4; Length 6;
Best Local Similarity 100.0%; Pred.No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
|||||
Db 1 HHHH 4

RESULT 88
US-09-115-397-1
; Sequence 1, Application US/09115397
; Patent No. 6576478
; GENERAL INFORMATION:
; APPLICANT: Wagner, Peter
; APPLICANT: Ault-Riche, Dana
; APPLICANT: NO. 6576478k, Steffen
; APPLICANT: Itin, Christian
; TITLE OF INVENTION: Microdevices for High-Throughput Screening of
; TITLE OF INVENTION: Biomolecules
; FILE REFERENCE: 24406-0002

CURRENT APPLICATION NUMBER: US/09/115.397
CURRENT FILING DATE: 1998-07-14
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 1

LENGTH: 6

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: histidine tag

US-09-115-397-1

Query Match 66.7%; Score 4; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5

||||

Db 1 HHHH 4

RESULT 89

US-09-463-194-2

Sequence 2, Application US/09463194

Patent No. 6585976

GENERAL INFORMATION:

APPLICANT: GICQUEL, BRIGITTE

APPLICANT: BERTHET, FRANCOIS-XAVIER

TITLE OF INVENTION: MYCOBACTERIUM STRAIN WITH MODIFIED ERP GENE AND VACCINE

FILE REFERENCE: 03715.0061-00000

CURRENT APPLICATION NUMBER: US/09/463.194

CURRENT FILING DATE: 2000-05-10

PRIOR APPLICATION NUMBER: PCT/FR98/01627

PRIOR FILING DATE: 1998-07-22

PRIOR APPLICATION NUMBER: FR 97/09303

PRIOR FILING DATE: 1997-07-22

NUMBER OF SEQ ID NOS: 5

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2

LENGTH: 6

TYPE: PRT

ORGANISM: Unknown Organism

FEATURE:

OTHER INFORMATION: Description of Unknown Organism: Illustrative

OTHER INFORMATION: C-terminal peptide

US-09-463-194-2

Query Match 66.7%; Score 4; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5

||||

Db 1 HHHH 4

RESULT 90

PCT-US94-04361-55

Sequence 55, Application PC/TUS9404361

GENERAL INFORMATION:

APPLICANT: Brigham and Women's Hospital

APPLICANT: 75 Francis Street

APPLICANT: Boston, MA 02115

APPLICANT: Bunn, H. Franklin

APPLICANT: Wen, Danyi

APPLICANT: Showers, Mark O.

TITLE OF INVENTION: Erythropoietin Muteins With Enhanced

TITLE OF INVENTION: Activity

NUMBER OF SEQUENCES: 59

CORRESPONDENCE ADDRESS:

ADDRESSEE: Steine, Kessler, Goldstein & Fox

STREET: 1100 New York Avenue, Suite 600

CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04361
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/049,802
FILING DATE: 21-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Cimballa, Michele A.
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0627.336PC01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: both
PCT-US94-04361-55

Query Match 66.7%; Score 4; DB 5; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5

||||

Db 1 HHHH 4

RESULT 91

US-08-276-852-14

Sequence 14, Application US/08276852

Patent No. 5652138

GENERAL INFORMATION:

APPLICANT: Burton, Dennis R

APPLICANT: Barbas, Carlos F

APPLICANT: Lerner, Richard A

TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES

TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS

NUMBER OF SEQUENCES: 170

CORRESPONDENCE ADDRESS:

ADDRESSEE: The Scripps Research Institute, Office of

ADDRESSEE: Patent Counsel

STREET: 10666 No. 5652138th Torrey Pines Road, Suite 220,

CITY: La Jolla

STATE: CA

COUNTRY: USA

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/276,852

FILING DATE: 18-JUL-1994

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/178,302

FILING DATE: 30-SEP-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/954,148

FILING DATE: 30-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCR1452P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: C-terminal
US-08-276-852-14

Query Match 66.7%; Score 4; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
Db 3 HHHH 6

RESULT 92
US-08-899-575-14
Sequence 14, Application US/08899575
Patent No. 5770440
GENERAL INFORMATION:
APPLICANT: Burton, Dennis R
APPLICANT: Barbas, Carlos F
APPLICANT: Lerner, Richard A
TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
NUMBER OF SEQUENCES: 170
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Scripps Research Institute, Office of
ADDRESSEE: Patent Counsel
STREET: 10666 No. 5770440th Torrey Pines Road, Suite 220,
STREET: Mail Drop TPC8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/899,575
FILING DATE: 24-JUL-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/276,852
FILING DATE: 18-JUL-1994
APPLICATION NUMBER: US 08/178,302
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/954,148
FILING DATE: 30-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCR1452P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: C-terminal
US-08-899-575-14

Query Match 66.7%; Score 4; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
Db 3 HHHH 6

RESULT 93
US-08-899-575-14
Sequence 14, Application US/08899575
Patent No. 5804440
GENERAL INFORMATION:
APPLICANT: Burton, Dennis R
APPLICANT: Barbas, Carlos F
APPLICANT: Lerner, Richard A
TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
NUMBER OF SEQUENCES: 170
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Scripps Research Institute, Office of
ADDRESSEE: Patent Counsel
STREET: 10666 No. 5804440th Torrey Pines Road, Suite 220,
STREET: Mail Drop TPC8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/899,575
FILING DATE: 24-JUL-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/276,852
FILING DATE: 18-JUL-1994
APPLICATION NUMBER: US 08/178,302
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/954,148
FILING DATE: 30-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCR1452P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: C-terminal
US-08-899-575-14

Query Match 66.7%; Score 4; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5

```
Db          3 HHHH 6

RESULT 94
JS-08-865-468-9
; Sequence 9, Application US/08865468
; Patent No. 6248869
; GENERAL INFORMATION:
; APPLICANT: Dade International Inc.
; APPLICANT: Morjana, Nihmat A.
; APPLICANT: Puia, Angela M.
; TITLE OF INVENTION: TROPONIN I FORMS AND USE OF SAME
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DADE INTERNATIONAL INC.
; STREET: 1717 Deerfield Road
; CITY: Deerfield
; STATE: Illinois
; COUNTRY: US
; ZIP: 60015
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/865,468
; FILING DATE: 29 May 1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: WINSTON, LOIS K.
; REGISTRATION NUMBER: 39,074
; REFERENCE/DOCKET NUMBER: DA-9018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (708) 267-5364
; TELEFAX: (708) 267-5376
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-865-468-9

Query Match 66.7%; Score 4; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 2 HHHH 5
Db 2 HHHH 5

RESULT 95
PCT-US95-08743-14
; Sequence 14, Application PC/TUS9508743
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/08743
; FILING DATE: 11-JUL-1995

Db          3 HHHH 6

PCT-US95-08743-14
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/276,852
; FILING DATE: 18-JUL-1994
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: C-terminal
; PCT-US95-08743-14

Query Match 66.7%; Score 4; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 2 HHHH 5
Db 3 HHHH 6

RESULT 96
US-08-762-433-6
; Sequence 6, Application US/08762433
; Patent No. 5877002
; GENERAL INFORMATION:
; APPLICANT: GRIFANTINI, RENATA
; APPLICANT: CARPANI, GIOVANNA
; APPLICANT: GALLI, GIULIANO
; APPLICANT: GRANDI, GUIDO
; TITLE OF INVENTION: THERMOSTABLE MUTANTS OF D-N-ALPHA-CARBAMOYLASE
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/762,433
; FILING DATE: 09-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IT MI95 A002700
; FILING DATE: 21-DEC-1995
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-762-433-6

Query Match 66.7%; Score 4; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 2 HHHH 5
Db 1 HHHH 4

RESULT 97
US-09-001-219-6
```

Sequence 6, Application US/09001219
Patent No. 5877003
GENERAL INFORMATION:
APPLICANT: GRIFANTINI, RENATA
APPLICANT: CARPANI, GIOVANNA
APPLICANT: GALLI, GIULIANO
APPLICANT: GRANDI, GUIDO
TITLE OF INVENTION: THERMOSTABLE MUTANTS OF D-N-ALPHA-CARBAMOYLASE
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: ORLON SPIVAK, MCCLELLAND, MAIER & NEUSTADT
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/001.219
FILING DATE: 21-DEC-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/762.433
FILING DATE: 09-DEC-1996
APPLICATION NUMBER: IT MI95 A002700
FILING DATE: 21-DEC-1995
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-09-001-219-6

Query Match 66.7%; Score 4; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
Db 1 HHHH 4

RESULT 98
US-08-481-658B-22
Sequence 22, Application US/08481658B
Patent No. 5955075
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/481,658B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3E
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-481-658B-22

Query Match 66.7%; Score 4; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
Db 3 HHHH 6

RESULT 99
US-08-477-504A-22
Sequence 22, Application US/08477504A
Patent No. 5972353
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,504A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-477-504A-22

Query Match 66.7%; Score 4; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
|||
Db 3 HHHH 6

RESULT 100
US-08-486-756A-22
; Sequence 22, Application US/08486756A
; Patent No. 5981711
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,756A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-486-756A-22

Query Match 66.7%; Score 4; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
|||
Db 3 HHHH 6

Search completed: October 21, 2003, 18:30:48
Job time : 21.6667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:24:06 ; Search time 19.3333 Seconds
(without alignments)
29.845 Million cell updates/sec

Title: US-10-057-890A-15
Perfect score: 6
Sequence: 1 GHRHHS 6

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1100

Minimum DB seq length: 0
Maximum DB seq length: 10

Post-processing: Listing first 150 summaries

Database : PIR_76:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3	50.0	9	2	phosphoenolpyruvat
2	2	33.3	3	3	growth-modulating
3	2	33.3	4	2	Carbon-monoxide de
4	2	33.3	5	2	COI intron 16 prot
5	2	33.3	5	2	biotin B - Citropa
6	2	33.3	5	2	hypothetical prote
7	2	33.3	6	2	sarcosine dehydrog
8	2	33.3	6	2	collagen alpha 1(I
9	2	33.3	6	2	Fc epsilon RIIB -
10	2	33.3	7	1	hypothalamic hepta
11	2	33.3	7	2	choline oxidase (E
12	2	33.3	7	2	gramicidin S synth
13	2	33.3	7	4	hypothetical L2 pr
14	2	33.3	8	2	urine glycopeptide
15	2	33.3	8	2	leucokinin VI - Ma
16	2	33.3	8	2	olfactory receptor
17	2	33.3	9	2	litorin 2-Glu - Au
18	2	33.3	9	2	litorin I - Austr
19	2	33.3	9	2	litorin - Rohde's
20	2	33.3	9	2	Ig heavy chain CDR
21	2	33.3	9	2	Ig heavy chain CDR
22	2	33.3	9	2	D-amino-acid oxida
23	2	33.3	10	1	leucosulfakinin-II
24	2	33.3	10	2	streptopain (EC 3.
25	2	33.3	10	2	neuromedin C - lau
26	2	33.3	10	2	neuromedin C - bov
27	2	33.3	10	2	erythrocyte membra
28	2	33.3	10	2	glucokinase (EC
29	2	33.3	10	2	peptide-N4-(N-acet

30	2	33.3	10	2	S30348	clotting protein -
31	2	33.3	10	2	B60656	leucosulfakinin II
32	2	33.3	10	2	A43977	PMRFamide-like pro
33	2	33.3	10	2	A30823	bothropstoxin - ja
34	2	33.3	10	2	A59173	nuclease Bhl (EC 3
35	1	16.7	3	3	RHTDIO	thyloliberin - Bom
36	1	16.7	3	3	RHPT	thyloliberin - pig
37	1	16.7	3	3	RHSHT	thyloliberin - she
38	1	16.7	3	3	A2971	thyloliberin - eas
39	1	16.7	3	3	A6898	bursin - chicken
40	1	16.7	3	3	A23751	spinal cord peptid
41	1	16.7	3	3	B23751	spinal cord peptid
42	1	16.7	3	3	I50412	gene p20K protein
43	1	16.7	3	3	PT0636	T-cell receptor be
44	1	16.7	3	3	PT0578	T-cell receptor be
45	1	16.7	3	3	PT0571	T-cell receptor be
46	1	16.7	3	3	PT0622	T-cell receptor be
47	1	16.7	4	1	ECXAA	antho-Ramide neur
48	1	16.7	4	2	A32039	tyrosine-melanocyt
49	1	16.7	4	2	A37832	phenol 2-monooxyge
50	1	16.7	4	2	A48360	gamma subunit of p
51	1	16.7	4	2	I40505	hypothetical prote
52	1	16.7	4	2	S17255	ribosomal protein
53	1	16.7	4	2	T30569	hypothetical prote
54	1	16.7	4	2	A25844	autho-Rf amide neu
55	1	16.7	4	2	A34626	RPCH-related neuro
56	1	16.7	4	2	A32480	achatin-I - giant
57	1	16.7	4	2	S39390	myosin-light chain
58	1	16.7	4	2	PT0240	Ig heavy chain CRD
59	1	16.7	4	2	PT0271	Ig mu chain V regi
60	1	16.7	4	2	S43959	T-cell receptor be
61	1	16.7	4	2	A53284	T-cell receptor be
62	1	16.7	4	2	B53284	T-cell receptor be
63	1	16.7	4	2	PT0696	T-cell receptor be
64	1	16.7	4	2	PT0645	T-cell receptor be
65	1	16.7	4	2	PT0633	T-cell receptor be
66	1	16.7	4	2	PT0711	T-cell receptor be
67	1	16.7	4	2	PT0661	T-cell receptor be
68	1	16.7	4	2	PT0712	T-cell receptor be
69	1	16.7	4	2	PT0534	T-cell receptor be
70	1	16.7	4	2	PT0698	T-cell receptor be
71	1	16.7	4	2	PT0677	T-cell receptor be
72	1	16.7	4	2	PT0551	T-cell receptor be
73	1	16.7	4	2	PT0697	T-cell receptor be
74	1	16.7	4	2	PT0706	T-cell receptor be
75	1	16.7	4	2	PT0675	T-cell receptor be
76	1	16.7	4	2	PT0721	T-cell receptor be
77	1	16.7	4	2	PT0566	T-cell receptor be
78	1	16.7	4	2	A40135	branched-chain-am
79	1	16.7	4	2	S47552	ubiquitin - rat
80	1	16.7	4	2	S09478	globulin IV alpha
81	1	16.7	5	2	A60521	glycogen phosphory
82	1	16.7	5	2	UN0860	peptidyl-dipeptida
83	1	16.7	5	2	A32516	cholecystokinin-5
84	1	16.7	5	2	C23751	spinal cord peptid
85	1	16.7	5	2	A26830	mitosis inhibiting
86	1	16.7	5	2	C41225	copper resistance
87	1	16.7	5	2	A41225	copper resistance
88	1	16.7	5	2	S70154	URP2 protein - Xan
89	1	16.7	5	2	I40702	primase - Citrobac
90	1	16.7	5	2	B37325	pap fibrillar regul
91	1	16.7	5	2	A44955	alkanal monooxygen
92	1	16.7	5	2	B31836	20K protein - Rick
93	1	16.7	5	2	I40469	dnazX-like protein
94	1	16.7	5	2	D60274	major protein anti
95	1	16.7	5	2	F22565	R-phycocerythrin ga
96	1	16.7	5	2	T14908	hypothetical prote
97	1	16.7	5	2	T14910	hypothetical prote
98	1	16.7	5	2	A33882	cadmium-binding pe
99	1	16.7	5	2	PQ0689	photosystem I 10.4
100	1	16.7	5	2	S51077	alpha-amylase - ri
101	1	16.7	5	2	B45525	actin I - malaria
102	1	16.7	5	2	A60803	neuropeptide - sea

103 1 16.7 5 2 A37114 hypoxanthine phosph
104 1 16.7 5 2 S65726 hemoglobin, extrac
105 1 16.7 5 2 B61445 Leu-enkephalin - b
106 1 16.7 5 2 A61445 Met-enkephalin - b
107 1 16.7 5 2 B61168 cocoonase (EC 3.4.
108 1 16.7 5 2 JH0253 gut pentapeptide -
109 1 16.7 5 2 S11075 alcohol dehydrogen
110 1 16.7 5 2 S53595 hypothetical prote
111 1 16.7 5 2 I50385 myosin light chain
112 1 16.7 5 2 S11127 phosphoprotein, bo
113 1 16.7 5 2 PT0267 ig heavy chain CRD
114 1 16.7 5 2 PT0278 ig heavy chain CRD
115 1 16.7 5 2 PT0281 ig heavy chain CRD
116 1 16.7 5 2 PT0295 ig heavy chain CRD
117 1 16.7 5 2 PT0308 ig heavy chain CRD
118 1 16.7 5 2 PT0520 ig kappa chain V-I
119 1 16.7 5 2 S62883 seminal plasma pro
120 1 16.7 5 2 D44823 synaptosomal-assoc
121 1 16.7 5 2 C53284 T-cell receptor be
122 1 16.7 5 2 PT0596 T-cell receptor be
123 1 16.7 5 2 PT0610 T-cell receptor be
124 1 16.7 5 2 PT0513 T-cell receptor be
125 1 16.7 5 2 PT0525 T-cell receptor be
126 1 16.7 5 2 PT0597 T-cell receptor be
127 1 16.7 5 2 PT0644 T-cell receptor be
128 1 16.7 5 2 PT0600 T-cell receptor be
129 1 16.7 5 2 PT0729 T-cell receptor be
130 1 16.7 5 2 PT0624 T-cell receptor be
131 1 16.7 5 2 PT0686 T-cell receptor be
132 1 16.7 5 2 PT0608 T-cell receptor be
133 1 16.7 5 2 PT0601 T-cell receptor be
134 1 16.7 5 2 PT0625 T-cell receptor be
135 1 16.7 5 2 PT0672 T-cell receptor be
136 1 16.7 5 2 PT0660 T-cell receptor be
137 1 16.7 5 2 PT0651 T-cell receptor be
138 1 16.7 5 2 PT0656 T-cell receptor be
139 1 16.7 5 2 PT0669 T-cell receptor be
140 1 16.7 5 2 PT0659 T-cell receptor be
141 1 16.7 5 2 PT0535 T-cell receptor be
142 1 16.7 5 2 PT0699 T-cell receptor be
143 1 16.7 5 2 PT0553 T-cell receptor be
144 1 16.7 5 2 PT0538 T-cell receptor be
145 1 16.7 5 2 PT0561 T-cell receptor be
146 1 16.7 5 2 PT0540 T-cell receptor be
147 1 16.7 5 2 PT0695 T-cell receptor be
148 1 16.7 5 2 PT0703 T-cell receptor be
149 1 16.7 5 2 PT0690 T-cell receptor be
150 1 16.7 5 2 PT0684 T-cell receptor be

ALIGNMENTS

RESULT 1
S13889
C:Species: Zea mays (maize)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Aug-1997
C:Accession: S13889
R:Jiao, J.; Chollet, R.
Arch. Biochem. Biophys., 283, 300-305, 1990
A:Title: Regulatory phosphorylation of serine-15 in maize phosphoenolpyruvate carboxylase
A:Reference number: S13889; MUID:91112741; PMID:2148863
A:Accession: S13889
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-9 <JIA>
C:Keywords: carbon-carbon lyase; carboxy-lyase

Query Match 50.0%; Score 3; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 HHS 6
Db 1 HHS 3

RESULT 2

GKHU

growth-modulating peptide - human

C:Species: Homo sapiens (man)

C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001

C:Accession: A01421

R:Schlesinger, D.H.; Pickart, L.; Thaler, M.M.

Experientia 33, 324-325, 1977

A:Title: Growth-modulating serum tripeptide is glycyl-histidyl-lysine.

A:Reference number: A01421; MUID:77162369; PMID:858356

A:Accession: A01421

A:Molecule type: protein

A:Residues: 1-3 <SCH>

A:Note: this serum tripeptide is found to stimulate growth of some cell types and to inh

C:Superfamily: unassigned animal peptides

Query Match: 33.3%; Score 2; DB 3; Length 3;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GH 2

Db 1 GH 2

RESULT 3

PL0140

carbon-monoxide dehydrogenase (EC 1.2.99.2) large chain - Pseudomonas carboxydohydrogena

C:Species: Pseudomonas carboxydohydrogena

C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 28-Apr-1993

C:Accession: PL0140

R:Kraut, M.; Hugendieck, I.; Herwig, S.; Meyer, O.

Arch. Microbiol. 152, 335-341, 1989

A:Title: Homology and distribution of CO dehydrogenase structural genes in carboxydotrop.

A:Reference number: PL0138; MUID:90055678; PMID:2818128

A:Accession: PL0140

A:Molecule type: protein

A:Residues: 1-4 <KRA>

C:Comment: Carbon-monoxide dehydrogenase consists of three polypeptide chains: large, me

C:Keywords: oxidoreductase

Query Match: 33.3%; Score 2; DB 2; Length 4;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GH 2

Db 2 GH 3

RESULT 4

I38888

COI intron 16 protein - Podospora anserina mitochondrion

C:Species: mitochondrion Podospora anserina

C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-Dec-1999

C:Accession: I38888

R:Cummings, D.J.; Michel, F.; McNally, K.L.

Curr. Genet. 16, 381-406, 1989

A:Title: DNA sequence analysis of the 24.5 kilobase pair cytochrome oxidase subunit I mi

A:Reference number: A48327; MUID:90124722; PMID:2558809

A:Accession: I38888

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-4 <CUM>

A:Cross-references: GB:X55026; GB:M30937; GB:M61734

C:Genetics:

A:Genome: mitochondrion

A:Genetic code: SGC3

C;Keywords: mitochondrion

Query Match 33.3%; Score 2; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GH 2
||
Db 3 GH 4

RESULT 5

140698
Biotin B - Citrobacter freundii (fragment)
C;Species: Citrobacter freundii
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 12-Aug-1996
C;Accession: I40698
R;Shuan, D.; Campbell, A.
Gene 67, 203-211, 1988
A;Title: Transcriptional regulation and gene arrangement of Escherichia coli, Citrobacter
A;Reference number: I40697; MUID:89006280; PMID:2971595
A;Accession: I40698
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-5 <RES>
A;Cross-references: GB:M21922; NID:gl44434

Query Match 33.3%; Score 2; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 HS 6
||
Db 3 HS 4

RESULT 6

T10954
Hypothetical protein 3 - spring vetch
C;Species: Vicia sativa (spring vetch, tare)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C;Accession: T10954
R;Christiansen, A.; Hansen, A.C.; Vijn, I.; Pallisgaard, N.; Larsen, K.; Yang, W.C.; Biss
submitted to the EMBL Data Library, December 1995
A;Description: A novel type of DNA binding protein interacts with a conserved sequence
A;Reference number: Z17228
A;Accession: T10954
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-5 <CHR>
A;Cross-references: EMBL:X95995; NID:gl360633; PID:e225862

Query Match 33.3%; Score 2; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 HS 6
||
Db 4 HS 5

RESULT 7

A61419
Sarcosine dehydrogenase (EC 1.5.99.1) - Pseudomonas sp. (strain WRF) (fragment)
C;Species: Pseudomonas sp.
C;Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 07-May-1999
C;Accession: A61419
R;Pinco, J.T.; Frisell, W.R.
Arch. Biochem. Biophys. 169, 483-491, 1975
A;Title: Characterization of the peptide-bound flavin of a bacterial sarcosine dehydroge
A;Reference number: A61419; MUID:76038634; PMID:241294
A;Accession: A61419
A;Molecule type: protein

A;Residues: 1-6 <PIN>

C;Keywords: PAD; flavoprotein; oxidoreductase; phosphoprotein
F;6/Modified site: 3'-PAD-histidine (His) #status experimental

Query Match 33.3%; Score 2; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GH 2
||
Db 5 GH 6

RESULT 8

B56979
collagen alpha 1(II) chain - bovine (fragment)
N;Alternate names: collagen alpha 3(XI) chain
C;Species: Bos primigenius taurus (cattle)
C;Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 03-Oct-1995
C;Accession: B56979
R;Wu, J.J.; Eyre, D.R.
J. Biol. Chem. 270, 18865-18870, 1995
A;Title: Structural analysis of cross-linking domains in cartilage type XI collagen.
A;Reference number: A56978; MUID:95370194; PMID:7642541
A;Accession: B56979
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-6 <WUA>
A;Note: the residue designated "X" is modified Lysine in collagen 1(II) some cross-li

Query Match 33.3%; Score 2; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GH 2
||
Db 4 GH 5

RESULT 9

A46474
Fc epsilon RIIB - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 11-Apr-1995
C;Accession: A46474
R;Richards, M.L.; Katz, D.H.; Liu, F.T.
J. Immunol. 147, 1067-1074, 1991
A;Title: Complete genomic sequence of the murine low affinity Fc receptor for IgE. Der
A;Reference number: A46474; MUID:91318149; PMID:1861070
A;Accession: A46474
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-6 <RIC>
A;Experimental source: BALB C, splenic B cells
A;Note: sequence extracted from NCBI backbone (NCBIP:45428)

Query Match 33.3%; Score 2; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HH 3
||
Db 4 HH 5

RESULT 10

NYPG7
hypothalamic heptapeptide - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 01-Sep-1981 #sequence_revision 01-Sep-1981 #text_change 23-Aug-1996
C;Accession: A01417
R;Chang, R.C.C.; Huang, W.Y.; Arimura, A.; Redding, T.W.; Coy, D.H.; Saifran, M.; Kone
Horm. Metab. Res. 13, 228-232, 1981

A;Title: Isolation, structure and synthesis of a heptapeptide with in vitro ACTH-release
 A;Reference number: A01417; MUID:81213980; PMID:6263778

A;Accession: A01417
 A;Molecule type: protein
 A;Residues: 1-7 <CHA>
 C;Superfamily: hypothalamic heptapeptide
 C;Keywords: hypothalamus

Query Match 33.3%; Score 2; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 HS 6

Db 4 HS 5

RESULT 11

Al5398
 choline oxidase (EC 1.1.3.17) - Alcaligenes sp. (tentative sequence) (fragment)

C;Species: Alcaligenes sp.
 C;Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 31-Mar-2000
 C;Accession: A15398
 R;Ohta-Fukuyama, M.; Miyake, Y.; Emi, S.; Yamano, T.
 J. Biochem. 88, 197-203, 1980

A;Title: Identification and properties of the prosthetic group of choline oxidase from *Alcaligenes*
 A;Reference number: A15398; MUID:81006769; PMID:6997283

A;Accession: A15398
 A;Molecule type: protein
 A;Residues: 1-7 <OHT>
 C;Keywords: oxidoreductase

Query Match 33.3%; Score 2; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 HS 6

Db 5 HS 6

RESULT 12

S42407
 gramicidin S synthetase component II - *Bacillus brevis* (fragment)

C;Species: *Bacillus brevis*
 C;Date: 20-Oct-1994 #sequence_revision 12-Apr-1996 #text_change 12-Apr-1996
 C;Accession: S42407

R;Kirstein, T.; Vater, J.; Kruff, J.; Wittmann-Liebold, B.; Franke, P.; Panico, M.; McDowell
 FEBS Lett. 340, 39-44, 1994
 A;Title: Detection of 4'-phosphopantetheine at the thioester binding site for L-valine
 A;Reference number: S42407; MUID:94164305; PMID:8119405

A;Accession: S42407
 A;Molecule type: protein
 A;Residues: 1-7 <STB>

Query Match 33.3%; Score 2; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GH 2

Db 3 GH 4

RESULT 13

I56695
 hypothetical L2 protein (mistranslated) - human papillomavirus type 16 (fragment)

C;Species: human papillomavirus type 16
 C;Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 18-Feb-2000
 C;Accession: I56695

R;Schneider-Maunoury, S.; Croissant, O.; Orth, G.
 J. Virol. 61, 3295-3298, 1987

A;Title: Integration of human papillomavirus type 16 DNA sequences: a possible early eve

A;Reference number: I56695; MUID:87311896; PMID:3041049

A;Accession: I56695
 A;Status: translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-7 <SCH>

A;Cross-references: GB:M30709; NID:gl90253; PIDN:AAA65995.1; PID:9553616
 C;Comment: This is the hypothetical translation of a viral sequence integrated into the
 C;Comment: It is translated in an incorrect, -1, reading frame of the L2 protein.

Query Match 33.3%; Score 2; DB 4; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 HS 6

Db 1 HS 2

RESULT 14

XGHEU

urine glycopeptide - human

C;Species: Homo sapiens (man)
 C;Date: 20-Jun-2000 #sequence_revision 20-Jun-2000 #text_change 20-Jun-2000
 C;Accession: A03188

R;Lote, C.J.; Weiss, J.B.
 Biochem. J. 123, 25P, 1971

A;Title: Identification in urine of a low-molecular-weight polar glycopeptide containing
 A;Reference number: A03188; MUID:72062338; PMID:5126885

A;Accession: A03188
 A;Molecule type: protein
 A;Residues: 1-8 <LOT>

C;Comment: The identity of the glycoprotein from which this peptide is derived is unknown.
 It has also been found (see PIR:XGHEU).

C;Superfamily: unassigned animal peptides

C;Keywords: glycoprotein

F;1/Binding site: carbohydrate (Cys) (covalent) #status experimental

Query Match 33.3%; Score 2; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 HS 6

Db 3 HS 4

RESULT 15

JS0316

leucokinin VI - Madeira cockroach

C;Species: *Leucophaea maderae* (Madeira cockroach)

C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Jun-2000

C;Accession: JS0316

R;Holman, G.M.; Cook, B.J.; Nachman, R.J.
 Comp. Biochem. Physiol. C 88, 27-30, 1987

A;Title: Isolation, primary structure, and synthesis of leucokinins V and VI: myotropic

A;Reference number: JS0315

A;Accession: JS0316

A;Molecule type: protein

A;Residues: 1-8 <HOL>

C;Comment: Leucokinins, a family of cephalomycotropic peptides, stimulate contractile act
 C;Keywords: amidated carboxyl end; cephalomycotropic peptide; pyroglutamic acid
 F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F;8/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 33.3%; Score 2; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 HS 6

Db 5 HS 6

```

RESULT 16
B54823
Olfactory receptor I7 - western wild mouse (fragment)
C:Species: Mus spretus (western wild mouse)
C>Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 17-Mar-1999
C:Accession: B54823
R:Chess, A.; Simon, I.; Cedar, H.; Axel, R.
Cell 78, 823-834, 1994
A>Title: Allelic inactivation regulates olfactory receptor gene expression.
A:Reference number: A54823; MUID:94373818; PMID:8087849
A:Accession: B54823
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-8 <CHE>

Query Match 33.3%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 HS 6
||
DB 6 HS 7

RESULT 17
S07205
litorin 2-Glu - Australian tree frog (Litoria aurea)
C:Species: Litoria aurea
C>Date: 12-Feb-1993 #sequence_revision 12-Mar-1993 #text_change 18-Aug-2000
C:Accession: S07205
R:Anastasi, A.; Montecucchi, P.; Angelucci, F.; Erspamer, V.; Endean, R.
Experientia 33, 1289, 1977
A>Title: Glu(OMe)(2)-litorin, the second bombesin-like peptide occurring in methanol ext
A:Reference number: S07205; MUID:78003546; PMID:908397
A:Accession: S07205
A:Molecule type: protein
A:Residues: 1-9 <ANA>
C:Superfamily: gastrin-releasing peptide
C:Keywords: amidated carboxyl end; neuropeptide; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:9/Modified site: amidated carboxyl end (Met) #status experimental

Query Match 33.3%; Score 2; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GH 2
||
DB 6 GH 7

RESULT 18
S07204
litorin I - Australian tree frog (Litoria aurea)
C:Species: Litoria aurea
C>Date: 12-Feb-1993 #sequence_revision 12-Mar-1993 #text_change 18-Aug-2000
C:Accession: S07204
R:Anastasi, A.; Erspamer, V.; Endean, R.
Experientia 31, 510-511, 1975
A>Title: Aminoacid composition and sequence of litorin, a bombesin-like nonapeptide from
A:Reference number: S07204; MUID:75187011; PMID:1140241
A:Accession: S07204
A:Molecule type: protein
A:Residues: 1-9 <ANA>
C:Superfamily: gastrin-releasing peptide
C:Keywords: amidated carboxyl end; neuropeptide; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:9/Modified site: amidated carboxyl end (Met) #status experimental

Query Match 33.3%; Score 2; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GH 2
||
DB 6 GH 7

RESULT 19
S07241
litorin - Rohde's leaf frog
C:Species: Phyllomedusa rohdei (Rohde's leaf frog)
C>Date: 12-Feb-1993 #sequence_revision 12-Mar-1993 #text_change 18-Aug-2000
C:Accession: S07241
R:Barra, D.; Falconieri Erspamer, G.; Simmaco, M.; Bossa, F.; Melchiorri, P.; Erspamer
FEBS Lett. 182, 53-56, 1985
A>Title: Rohdei-litorin: a new peptide from the skin of Phyllomedusa rohdei.
A:Reference number: S07241; MUID:85127560; PMID:3838283
A:Accession: S07241
A:Molecule type: protein
A:Residues: 1-9 <BAR>
C:Superfamily: gastrin-releasing peptide
C:Keywords: amidated carboxyl end; blocked amino end; neuropeptide; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:9/Modified site: amidated carboxyl end (Met) #status experimental

Query Match 33.3%; Score 2; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GH 2
||
DB 6 GH 7

RESULT 20
PT0231
Ig heavy chain CDR3 region (clone 1-118B) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C:Accession: PT0231
R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A>Title: Preferential utilization of specific immunoglobulin heavy chain diversity an
A:Reference number: PT0222; MUID:91108337; PMID:1899102
A:Accession: PT0231
A:Molecule type: DNA
A:Residues: 1-9 <YAM>
A:Experimental source: B lymphocyte
C:Keywords: heterotetramer; immunoglobulin

Query Match 33.3%; Score 2; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 HS 6
||
DB 3 HS 4

RESULT 21
PT0238
Ig heavy chain CDR3 region (clone 2-94B) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C:Accession: PT0238
R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A>Title: Preferential utilization of specific immunoglobulin heavy chain diversity an
A:Reference number: PT0222; MUID:91108337; PMID:1899102
A:Accession: PT0238
A:Molecule type: DNA
A:Residues: 1-9 <YAM>
A:Experimental source: B lymphocyte
C:Keywords: heterotetramer; immunoglobulin

```

Query Match 33.3%; Score 2; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 HS 6
 ||
 Db 6 HS 7

RESULT 22

Q-aminic-acid oxidase (EC 1.4.3.3) - Trigonopsis variabilis (fragment)

C;Species: Trigonopsis variabilis
 C;Date: 19-Mar-1997 #sequence_revision 05-Dec-1997 #text_change 07-May-1999
 C;Accession: S39437

R;Conlon, J.R.; Schraeder, T.; Andreesen, J.R.

A;Title: Evidence for the functional importance of Cys298 in D-amino acid oxidase from

A;Reference number: S39437; MUID:94094869; PMID:7903639

A;Accession: S39437

A;Molecule type: protein

A;Residues: 1-9 <SCH>

A;Experimental source: CBS 4095

A;Function:

C;Description: oxidoreductase; catalyzes the oxidation of D-amino acids to their correspond-
 ing alpha-keto acids; Note: reoxidation of the enzyme by molecular oxygen is accompanied by the release of H₂O;
 C;Keywords: FAD; oxidoreductase

Query Match 33.3%; Score 2; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GH 2
 ||
 Db 3 GH 4

RESULT 23

Q-aminic-acid oxidase (EC 1.4.3.3) - Madeira cockroach

C;Species: Leucophaea maderae (Madeira cockroach)

C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 13-Sep-1996

C;Accession: A26335

R;Conlon, J.R.; Schraeder, T.; Andreesen, J.R.

A;Title: Evidence for the functional importance of Cys298 in D-amino acid oxidase from

A;Reference number: A26335; MUID:87048769; PMID:3778455

A;Accession: A26335

A;Molecule type: protein

A;Residues: 1-10 <NAC>

A;Comment: This peptide was isolated from head extracts. It stimulates muscle contractio

C;Superfamily: gastrin

C;Keywords: amidated carboxyl end; neuropeptide; pyroglutamic acid; sulfoprotein

F;10/Modified site: pyroglutamic acid (Gln) #status experimental

F;5/Binding site: sulfate (Tyr) (covalent) #status experimental

F;10/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 33.3%; Score 2; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 3.4e+03;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GH 2

||

Db 6 GH 7

RESULT 24

A61289

C;Species: Streptococcus sp.

C;Date: 27-Jun-1994 #sequence_revision 27-Jun-1994 #text_change 07-May-1999

C;Accession: A61289

A;Title: Evidence for the functional importance of Cys298 in D-amino acid oxidase from

A;Reference number: A61289; MUID:94094869; PMID:7903639

A;Accession: A61289

A;Molecule type: protein

A;Residues: 1-10 <NAC>

A;Comment: This peptide was isolated from head extracts. It stimulates muscle contractio

R;Lo, S.S.; Fraser, B.A.; Liu, T.Y.

A;Title: The mixed disulfide in the zymogen of streptococcal proteinase. Characterization

A;Reference number: A61289; MUID:84289536; PMID:8381494

A;Accession: A61289

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-10 <LOA>

C;Superfamily: streptococcal cysteine proteinase

C;Keywords: cysteine proteinase; hydrolase

Query Match 33.3%; Score 2; DB 2; Length 10;

Best Local Similarity 100.0%; Pred. No. 3.4e+03;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GH 2

||

Db 7 GH 8

RESULT 25

PQ0177

neurotrophin C - laughing frog

C;Species: Rana ridibunda (laughing frog)

C;Date: 23-Nov-1991 #sequence_revision 23-Nov-1991 #text_change 11-Jan-2000

C;Accession: PQ0177

R;Conlon, J.M.; O'Harte, F.; Vaudry, H.

A;Title: Primary structures of the bombesin-like neuropeptides in frog brain show that b

A;Reference number: PQ0177; MUID:91315477; PMID:1859413

A;Accession: PQ0177

A;Molecule type: protein

A;Residues: 1-10 <CON>

A;Experimental source: brain

C;Superfamily: gastrin-releasing peptide

C;Keywords: amidated carboxyl end

F;10/Modified site: amidated carboxyl end (Met) #status predicted

Query Match 33.3%; Score 2; DB 2; Length 10;

Best Local Similarity 100.0%; Pred. No. 3.4e+03;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GH 2

||

Db 7 GH 8

RESULT 26

A60647

neurotrophin C - bovine

C;Species: Bos primigenius taurus (cattle)

C;Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 07-May-1999

C;Accession: A60647

R;Lemaire, S.; Trifaro, J.M.; Chouinard, L.; Cecyre, D.; Dessureault, J.; Mercier, P.; D

Peptides 10, 355-360, 1989

A;Title: Structural identification, subcellular localization and secretion of bovine adr

A;Reference number: A60647; MUID:89331342; PMID:2755876

A;Accession: A60647

A;Molecule type: protein

A;Residues: 1-10 <LEM>

A;Note: this neuropeptide was purified from secretory granules of cells in the adrenal m

C;Superfamily: gastrin-releasing peptide

C;Keywords: adrenal gland; neuropeptide

Query Match 33.3%; Score 2; DB 2; Length 10;

Best Local Similarity 100.0%; Pred. No. 3.4e+03;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GH 2

||

Db 7 GH 8

```

RESULT 27
XGHUE
erythrocyte membrane glycopeptide - human
C:Species: Homo sapiens (man)
C:Date: 20-Jun-2000 #sequence_revision 20-Jun-2000 #text_change 20-Jun-2000
C:Accession: A03187
R:Weiss, J.B.; Lote, C.J.; Bobinski, H.
Nature New Biol. 234, 25-26, 1971
A:Title: New low molecular weight glycopeptide containing triglycosylcysteine in human
A:Accession: A03187
A:Molecule type: protein
A:Residues: 1-10 <WEI>
C:Comment: The identity of the glycoprotein from which this peptide is derived is unknown
C:Superfamily: unassigned animal peptides
C:Keywords: glycoprotein
F:1/Binding site: carbohydrate (Cys) (covalent) #status experimental

Query Match 33.3%; Score 2; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.4e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GH 2
Db 3 GH 4

RESULT 28
S74176
glucokinase (EC 2.7.1.12), thermoresistant - Escherichia coli (fragment)
C:Species: Escherichia coli
C:Date: 14-Apr-1998 #sequence_revision 24-Apr-1998 #text_change 07-May-1999
C:Accession: S74176
R:Izu, H.; Adachi, O.; Yamada, M.
FEBS Lett. 394, 14-16, 1996
A:Title: Purification and characterization of the Escherichia coli thermoresistant gluc
A:Reference number: S74176; MUID:97074194; PMID:8925917
A:Accession: S74176
A:Molecule type: protein
A:Residues: 1-10 <IZU>
A:Experimental source: strain K-12
C:Genetics:
A:Gene: gntK
C:Keywords: dimer; phosphotransferase

Query Match 33.3%; Score 2; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.4e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HH 3
Db 7 HH 8

RESULT 29
A59272
peptide-N4-(N-acetyl-beta-glucosaminyl)asparagine amidase (EC 3.5.1.52) A, large chain
N:Alternate names: peptide N-glycosidase
C:Species: Prunus dulcis var. sativa (sweet almond)
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-May-2000
C:Accession: A59272
R:Altman, F.; Paschinger, K.; Dalik, T.; Vorauer, K.
Eur. J. Biochem. 252, 118-123, 1998
A:Title: Characterization of peptide-N4-(N-acetyl-beta-glucosaminyl)asparagine amidase
A:Reference number: A59272; MUID:98181894; PMID:9523720
A:Accession: A59272
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <ALT>
C:Keywords: hydrolase

Query Match 33.3%; Score 2; DB 2; Length 10;

Best Local Similarity 100.0%; Pred. No. 3.4e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GH 2
Db 6 GH 7

RESULT 32
A43977
FMRamide-like protein - tobacco hornworm
C:Species: Manduca sexta (tobacco hornworm)
C:Date: 11-Feb-1993 #sequence_revision 11-Feb-1993 #text_change 17-Mar-1999
C:Accession: A43977
R:Kings, T.G.; Teplov, D.B.; Phillips, J.M.; Riehm, J.P.; Rao, K.R.; Hildebrand, J.G.
Peptides 11, 849-856, 1990
A:Title: A new peptide in the FMRamide family isolated from the CNS of the hawkmoth,
A:Reference number: A43977; MUID:91045350; PMID:2235684
A:Accession: A43977
A:Molecule type: protein
A:Residues: 1-10 <KIN>
C:Keywords: amidated carboxyl end; neuroptide; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Best Local Similarity 100.0%; Pred. No. 3.4e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 HS 6
Db 6 HS 7

RESULT 30
S30348
clotting protein - signal crayfish
C:Species: Pacifastacus leniusculus (signal crayfish)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 17-Mar-1999
C:Accession: S30348
R:Kopacek, P.; Hall, M.; Soederhaell, K.
Eur. J. Biochem. 213, 591-597, 1993
A:Title: Characterization of a clotting protein, isolated from plasma of the freshwater
A:Reference number: S30348; MUID:93238739; PMID:8097463
A:Accession: S30348
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <KOP>

Query Match 33.3%; Score 2; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.4e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 HS 6
Db 2 HS 3

RESULT 31
B60656
leucosulfakinin II, non-sulfated - American cockroach
C:Species: Periplaneta americana (American cockroach)
C:Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 08-Dec-1995
C:Accession: B60656
R:Veenstra, J.A.
Neuropeptides 14, 145-149, 1989
A:Title: Isolation and structure of two gastrin/CKK-like neuropeptides from the Americ
A:Reference number: A60656; MUID:90137190; PMID:2615921
A:Accession: B60656
A:Molecule type: protein
A:Residues: 1-10 <VEE>
C:Keywords: amidated carboxyl end; neuroptide; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 33.3%; Score 2; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.4e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GH 2
Db 6 GH 7

```

F;10/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 33.3%; Score 2; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.4e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 HS 6
||
Db 5 HS 6

RESULT 33
R30823
bothropstoxin - jararacaeu (fragment)
C:Species: Bothrops jararacaeu (Jararacaeu)
C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 18-Jun-1993
C:Accession: A30823
Toxicon 26, 615-627, 1988
A:Title: Fractionation of Bothrops jararacaeu snake venom: partial chemical characteriz
A:Reference number: A30823; MUID:89020120; PMID:3176051
A:Accession: A30823
A:Molecule type: protein
A:Residues: 1-10 <HOM>

Query Match 33.3%; Score 2; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.4e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GH 2
||
Db 6 GH 7

RESULT 34
AS9173
nuclease Bhl (EC 3.1.1.-) - Basidiobolus haptosporus (fragment)
C:Species: Basidiobolus haptosporus
C:Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 18-Feb-2000
C:Accession: AS9173
submitted to the Protein Sequence Database, February 2000
A:Description: Single-strand-specific, guanylic acid preferential nuclease from Basidiob
A:Reference number: AS9173
A:Accession: AS9173
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <DES>
A:Note: extracellular, single-strand-specific nuclease
C:Keywords: hydrolase

Query Match 33.3%; Score 2; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.4e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GH 2
||
Db 5 GH 6

RESULT 35
RHTD0
thyroliberin - Bombina orientalis
C:Species: Bombina orientalis
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: A90919; A01415
R:Yasuhara, T.; Nakajima, T.
Chem. Pharm. Bull. 23, 3301-3303, 1975
A:Title: Occurrence of Pyr-His-Pro-NH-2 in the frog skin.
A:Reference number: A90919; MUID:76138399; PMID:815011
A:Accession: A90919
A:Molecule type: protein
A:Residues: 1-3 <YAS>

C:Superfamily: thyroliberin precursor
C:Keywords: amidated carboxyl end; cutaneous gland; hormone; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 16.7%; Score 1; DB 3; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 H 2
|
Db 2 H 2

RESULT 36
RHPGT
thyroliberin - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: A01415
R:Rair, R.M.G.; Barrett, J.F.; Bowers, C.Y.; Schally, A.V.
Biochemistry 9, 1103-1106, 1970
A:Title: Structure of porcine thyrotropin releasing hormone.
A:Reference number: A90560; MUID:70136150; PMID:4984938
A:Accession: A01415
A:Molecule type: protein
A:Residues: 1-3 <NAL>
R:Boiler, J.; Enmann, F.; Folkers, K.; Bowers, C.Y.; Schally, A.V.
Biochem. Biophys. Res. Commun. 37, 705-710, 1969
A:Title: The identity of chemical and hormonal properties of the thyrotropin releasing h
A:Reference number: A90167; MUID:70039904; PMID:4982117
A:Contents: annotation
A:Note: biological activities and Rf values (in 17 chromatographic systems) of the synth
C:Superfamily: thyroliberin precursor
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 16.7%; Score 1; DB 3; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 H 2
|
Db 2 H 2

RESULT 37
RHSHT
thyroliberin - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: A93750; A01415
R:Desiderio Jr., D.W.; Burgess, R.; Dunn, T.F.; Vale, W.; Guillemin, R.; Ward, D.N.
Org. Mass Spectrom. 5, 221-228, 1971
A:Title: The elucidation of the primary structure of the hypothalamic thyroid stimulat
A:Reference number: A93750
A:Accession: A93750
A:Molecule type: protein
A:Residues: 1-3 <DES>
R:Burgus, R.; Dunn, T.F.; Desiderio, D.; Ward, D.N.; Vale, W.; Guillemin, R.
Nature 226, 321-325, 1970
A:Title: Characterization of ovine hypothalamic hypophysiotropic TSH-releasing factor.
A:Reference number: A93161; MUID:70163386; PMID:4985794
A:Contents: annotation
A:Note: physicochemical characteristics and biological activities of the natural and syn
C:Superfamily: thyroliberin precursor
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 16.7%; Score 1; DB 3; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 H 2

Db 2 H 2

RESULT 38

A29271
thyroliberin - eastern newt (tentative sequence)
C:Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)
C>Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: A92971; A01415
R:Grimm-Jorgensen, Y.; McKelvy, J.F.
J. Neurochem. 23, 471-478, 1974

A:Title: Biosynthesis of thyrotropin releasing factor by newt (Triturus viridescens) brain
A:Reference number: A92971; MUID:75035605; PMID:4214528
A:Accession: A92971

A:Molecule type: protein

A:Residues: 1-3 <GRI>

A>Note: a peptide with the chromatographic and electrophoretic characteristics of thyroliberin, or glutamic acid

C:Superfamily: thyroliberin precursor

C:Keywords: amidated carboxyl end; cutaneous gland; hormone; hypothalamus; pyroglutamic

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 16.7%; Score 1; DB 3; Length 3;

Best Local Similarity 100.0%; Pred.No. 2.8e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 H 2

Db 2 H 2

RESULT 39

A60898
bursin - chicken
C:Species: Gallus gallus (chicken)
C>Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: A60898

R:Audhya, T.; Kroon, D.; Heavner, G.; Viamontes, G.; Goldstein, G.

Science 231, 997-999, 1986

A:Title: Tripeptide structure of bursin, a selective B-cell-differentiating hormone of the chicken

A:Reference number: A60898; MUID:86122916; PMID:3484838

A:Accession: A60898

A:Molecule type: protein

A:Residues: 1-3 <AUD>

C:Superfamily: unassigned animal peptides

C:Keywords: amidated carboxyl end; hormone

F:3/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 16.7%; Score 1; DB 3; Length 3;

Best Local Similarity 100.0%; Pred.No. 2.8e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 H 2

Db 2 H 2

RESULT 40

A23751
spinal cord peptide SCP-4 - pig
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: A23751

R:Hei, K.L.; Chen, R.L.; Zhang, H.L.; Lu, Y.A.; Guo, S.Y.; Wu, S.X.; Tsou, K.

Arch. Biochem. Biophys. 240, 178-183, 1985

A:Reference number: A23751; MUID:85250425; PMID:4015098

A:Accession: A23751

A>Status: preliminary

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions

A:Molecule type: protein
A:Residues: 1-3 <HSI>
C:Superfamily: unassigned animal peptides

Query Match 16.7%; Score 1; DB 3; Length 3;

Best Local Similarity 100.0%; Pred.No. 2.8e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1

Db 2 G 2

RESULT 41

B23751
spinal cord peptide SCP-5 - pig

C:Species: Sus scrofa domestica (domestic pig)

C>Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001

C:Accession: B23751

R:Hei, K.L.; Chen, R.L.; Zhang, H.L.; Lu, Y.A.; Guo, S.Y.; Wu, S.X.; Tsou, K.

Arch. Biochem. Biophys. 240, 178-183, 1985

A:Reference number: A23751; MUID:85250425; PMID:4015098

A:Accession: B23751

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-3 <HSI>

C:Superfamily: unassigned animal peptides

Query Match 16.7%; Score 1; DB 3; Length 3;

Best Local Similarity 100.0%; Pred.No. 2.8e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1

Db 3 G 3

RESULT 42

I50412
gene p20K protein - chicken (fragment)

C:Species: Gallus gallus (chicken)

C>Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001

C:Accession: I50412

R:Mao, P.L.; Beauchemin, M.; Bedard, P.A.

J. Biol. Chem. 268, 8131-8139, 1993

A:Title: Quiescence-dependent activation of the p20K promoter in growth-arrested chick

A:Reference number: A46643; MUID:93216790; PMID:8463325

A:Accession: I50412

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-3 <MAO>

A:Cross-references: GB:I02537; NID:G212616; PID:G212617

C:Genetics:

A:Gene: p20K

Query Match 16.7%; Score 1; DB 3; Length 3;

Best Local Similarity 100.0%; Pred.No. 2.8e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6

Db 2 S 2

RESULT 43

PT0636
T-cell receptor beta chain V-D-J region (100-2AT) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001

C:Accession: PT0636

R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions

C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001

C;Accession: PT0509; PMID:1711558

R;Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A;Molecule type: mRNA

A;Residues: 1-3 <FEE>

A;Experimental source: newborn thymus, strain BALB/c

C;Keywords: T-cell receptor

Query Match 16.7%; Score 1; DB 3; Length 3;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 G 1

Db 3 G 3

RESULT 44

PT0578

T-cell receptor beta chain V-D-J region (141-LBD) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001

C;Accession: PT0578

R;Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A;Molecule type: mRNA

A;Residues: 1-3 <FEE>

A;Experimental source: day 19 fetal thymus, strain BALB/c

C;Keywords: T-cell receptor

Query Match 16.7%; Score 1; DB 3; Length 3;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6

Db 2 S 2

RESULT 45

PT0571

T-cell receptor beta chain V-D-J region (141-LCM) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001

C;Accession: PT0571

R;Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A;Molecule type: mRNA

A;Residues: 1-3 <FEE>

A;Experimental source: day 19 fetal thymus, strain BALB/c

C;Keywords: T-cell receptor

Query Match 16.7%; Score 1; DB 3; Length 3;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6

Db 2 S 2

RESULT 46

PT0622

T-cell receptor beta chain V-D-J region - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001

C;Accession: PT0622; PT0680; PT0582; PT0673

R;Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A;Reference number: PT0509; PMID:1711558

A;Accession: PT0622

A;Status: translation not shown

A;Molecule type: mRNA

A;Residues: 1-3 <FEE>

A;Experimental source: newborn thymus, strain BALB/c, clone 111-1P

A;Accession: PT0680

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-3 <FEE1>

A;Experimental source: day 18 fetal thymus, strain BALB/c clones 154-1J and 154-1F

A;Accession: PT0582

A;Status: translation not shown

A;Molecule type: mRNA

A;Residues: 1-3 <FEE2>

A;Experimental source: day 19 fetal thymus, strain BALB/c, clone 159-2A

C;Keywords: T-cell receptor

Query Match 16.7%; Score 1; DB 3; Length 3;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6

Db 2 S 2

RESULT 47

ECXAA

antho-Branide neuropeptide - sea anemone (Anthopleura elegantissima)

C;Species: Anthopleura elegantissima

C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 08-Dec-1995

C;Accession: A26666

R;Grimmelikhuijzen, C.J.P.; Graff, D.

Proc. Natl. Acad. Sci. U.S.A. 83, 9817-9821, 1986

A;Title: Isolation of <Glu-Gly-Arg-Phe-NH2 (Antho-RFamide), a neuropeptide from sea anem

A;Reference number: A26666; PMID:87092339; PMID:2879288

A;Accession: A26666

A;Molecule type: protein

A;Residues: 1-4 <GR1>

C;Comment: The function of this peptide is not known but it could act as a transmitter a

C;Superfamily: RFamide neuropeptide

C;Keywords: amidated carboxyl end; neuropeptide; pyroglutamic acid

F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F;4/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 16.7%; Score 1; DB 1; Length 4;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 G 1

Db 2 G 2

RESULT 48

A32039

tyrosine-melanocyte-stimulating hormone release-inhibiting factor 1 - bovine

C;Species: Bos primigenius taurus (cattle)

C;Date: 31-Jul-1989 #sequence_revision 31-Jul-1989 #text_change 18-Aug-2000

C;Accession: A32039

R;Horvath, A.; Kastin, A.J.

J. Biol. Chem. 264, 2175-2179, 1989

A;Title: Isolation of tyrosine-melanocyte-stimulating hormone release-inhibiting factor

A;Reference number: A32039; PMID:89123285; PMID:2563371

A;Accession: A32039

A;Molecule type: protein

A;Residues: 1-4 <HOR>

A;Experimental source: brain

C;Superfamily: unassigned animal peptides

C;Keywords: amidated carboxyl end

F;4/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 16.7%; Score 1; DB 2; Length 4;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 G 1

Db 4 G 4

RESULT 49

A37832 phenol 2-monooxygenase (EC 1.14.13.7) chain P5 - Pseudomonas sp. (strain CP600) (fragment)

C;Species: Pseudomonas sp.

C;Date: 14-Jun-1991 #sequence_revision 14-Jun-1991 #text_change 23-Jun-1993

C;Accession: A37832

R;Powlowski, J.; Shingler, V.

J. Bacteriol. 172, 6834-6840, 1990

A;Title: In vitro analysis of polypeptide requirements of multicomponent phenol hydroxylase

A;Reference number: A37832; MUID:91072231; PMID:2254259

A;Accession: A37832

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-4 <POW>

C;Keywords: oxidoreductase

Query Match

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6

Db 1 S 1

RESULT 50

A48360 gamma subunit of Protein A - Methylosinus trichosporium (fragment)

C;Species: Methylosinus trichosporium

C;Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 08-Oct-1999

C;Accession: A48360

R;Cardy, D.L.; Laidler, V.; Salmond, G.P.; Murrell, J.C.

Arch. Microbiol. 156, 477-483, 1991

A;Title: The methane monooxygenase gene cluster of Methylosinus trichosporium: cloning and

A;Reference number: A48360; MUID:92153031; PMID:1785954

A;Contents: OB3b

A;Accession: A48360

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-4 <CAR>

A;Cross-references: GB:S81887; NID:g245213; PIDN:AAB321391.1; PID:g245214

A;Note: sequence extracted from NCBI backbone (NCBIN:81887, NCBIP:81912)

Query Match

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 H 2

Db 1 H 1

RESULT 51

I40505

hypothetical protein 3 (4 aa) - Bacillus stearothermophilus

C;Species: Bacillus stearothermophilus

C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 15-Oct-1999

C;Accession: I40505

R;Waye, M.M.; Winter, G.

Eur. J. Biochem. 159, 505-510, 1986

A;Title: A transcription terminator in the 5' non-coding region of the tyrosyl tRNA s

A;Reference number: I40503; MUID:86274732; PMID:3525162

A;Accession: I40505

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-4 <RSS>

A;Cross-references: EMBL:X04193; NID:g40233; PIDN:CAA27783.1; PID:g580944

Query Match

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6

Db 3 S 3

RESULT 52

S17255

ribosomal protein YmL1, mitochondrial, questionable - yeast (Saccharomyces cerevisiae)

C;Species: Saccharomyces cerevisiae

A;Variety: strain 07173

C;Date: 23-Apr-1993 #sequence_revision 14-Sep-1994 #text_change 09-May-1997

C;Accession: S17255

R;Grohmann, L.; Graack, H.R.; Kruft, V.; Choli, T.; Goldschmidt-Reisin, S.; Kitakawa,

FEBS Lett. 284, 51-56, 1991

A;Title: Extended N-terminal sequencing of proteins of the large ribosomal subunit fr

A;Reference number: S17255; MUID:91285106; PMID:2060626

A;Accession: S17255

A;Molecule type: protein

A;Residues: 1-4 <GRO>

C;Comment: A coding region for this protein could not be identified in the genome of :

C;Genetics:

A;Genome: nuclear

C;Keywords: mitochondrion; protein biosynthesis; ribosome

Query Match

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6

Db 1 S 1

RESULT 53

T30569

hypothetical protein - Emericella nidulans

C;Species: Emericella nidulans, Aspergillus nidulans

C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000

C;Accession: T30569

R;Morrice, J.; MacKenzie, D.A.; Parr, A.J.; Archer, D.B.

Curr. Genet. 34, 379-385, 1998

A;Title: Isolation and characterisation of the acetyl-CoA carboxylase gene from Asperg

A;Reference number: Z20869; MUID:99087906; PMID:9871120

A;Accession: T30569

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-4 <MOR>

A;Cross-references: EMBL:Y15996; NID:e1285512; PID:e1218041; PIDN:CAA75927.1

Query Match

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 G 1

Db 2 G 2

RESULT 54

```

A25844
autotho-RF amide neuropeptide - sea pansy (Renilla koellikeri)
C;Species: Renilla koellikeri (Koelliker's sea pansy)
C;Date: 21-May-1988 #sequence_revision 30-Sep-1993 #text_change 11-Jul-1997
C;Accession: A25844
FEBS Lett. 211, 105-108, 1987
A;Title: Isolation of the neuropeptide pGlu-Gly-Arg-Phe-amide from the pennatulid Renilla
A;Reference number: A25844
A;Accession: A25844
A;Molecule type: protein
A;Residues: 1-4 <GRI>
C;Keywords: amidated carboxyl end; neuropeptide; pyroglutamic acid
F;2/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
Query Match 16.7%; Score 1; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 G 1
Db 1 G 1

RESULT 55
A25844
RRCH-related neuropeptide - ferruginous spindle
C;Species: Fuscus ferrugineus (ferruginous spindle)
C;Date: 06-Jul-1990 #sequence_revision 06-Jul-1990 #text_change 31-Dec-1993
C;Accession: A34626
Biochem. Biophys. Res. Commun. 167, 273-279, 1990
A;Title: A molluscan neuropeptide related to the crustacean hormone, RRCH.
A;Reference number: A34626; MUID:90179762; PMID:2310394
A;Accession: A34626
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-4 <KUR>
C;Keywords: neuropeptide

Query Match 16.7%; Score 1; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 G 1
Db 2 G 2

RESULT 56
A32480
Achatin-I - giant African snail
N;Contents: achatin-II
C;Species: Achatina fulica (giant African snail)
C;Date: 12-Oct-1989 #sequence_revision 12-Oct-1989 #text_change 17-Mar-1999
C;Accession: A32480
Biochem. Biophys. Res. Commun. 160, 1015-1020, 1989
A;Title: Achatin-I, an endogenous neuroexcitatory tetrapeptide from Achatina fulica feru
A;Reference number: A32480; MUID:89273551; PMID:2597281
A;Accession: A32480
A;Molecule type: protein
A;Residues: 1-4 <KAM>
A;Note: stereochemistry of the active form confirmed by chemical synthesis
R;Rishida, T.; In, Y.; Inoue, M.; Yasuda-Kamatani, Y.; Minakata, H.; Iwashita, T.; Nomoto
FEBS Lett. 307, 253-256, 1992
A;Title: Effect of the D-Phe(2) residue on molecular conformation of an endogenous neuro
(H-Gly-Phe-Ala-Asp-OH).
A;Reference number: A44691; MUID:92354723; PMID:1644179
A;Contents: annotation; X-ray crystallography, 0.85 angstroms
A;Note: achatin-II has L-phenylalanine
C;Keywords: D-amino acid

F;2/Modified site: D-phenylalanine (Phe) #status experimental
Query Match 16.7%; Score 1; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 G 1
Db 1 G 1

RESULT 57
S39390
myosin-light-chain kinase (EC 2.7.1.117), smooth muscle - turkey (fragment)
C;Species: Meleagris gallopavo (common turkey)
C;Date: 19-Mar-1997 #sequence_revision 10-Oct-1997 #text_change 07-May-1999
C;Accession: S39390
R;Komatsu, H.; Ikebe, M.
Biochem. J. 296, 53-58, 1993
A;Title: Affinity labelling of smooth-muscle myosin light-chain kinase with 5'-(p-(fluor
A;Reference number: S39390; MUID:94071841; PMID:8250857
A;Accession: S39390
A;Molecule type: protein
A;Residues: 1-4 <KOW>
A;Experimental source: gizzard
C;Keywords: phosphotransferase; smooth muscle

Query Match 16.7%; Score 1; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 G 1
Db 2 G 2

RESULT 58
PT0240
Ig heavy chain CRD3 region (clone 2-100B) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0240
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and j
A;Reference number: PT0222; MUID:91108337; PMID:1899102
A;Accession: PT0240
A;Molecule type: DNA
A;Residues: 1-4 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin

Query Match 16.7%; Score 1; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 G 1
Db 3 G 3

RESULT 59
PT0271
Ig heavy chain CRD3 region (clone 3-103A) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0271
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and j
A;Reference number: PT0222; MUID:91108337; PMID:1899102
A;Accession: PT0271
A;Molecule type: DNA

```

A;Residues: 1-4 <YAM>
 A;Experimental source: B lymphocyte
 C;Keywords: heterotetramer; immunoglobulin

Query Match 16.7%; Score 1; DB 2; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 Db 2 G 2

RESULT 60

S43959
 IG mu chain V region (clone l3) - human (fragment)
 C;Species: Homo sapiens (man)
 C;Date: 20-Oct-1994 #sequence_revision 26-Jul-1996 #text_change 17-Mar-1999
 C;Accession: S43959
 R;Wagner, S.D.; Williams, G.T.; Larson, T.; Neuberger, M.S.; Kitamura, D.; Rajewsky, K.;
 Nucleic Acids Res. 22, 1389-1393, 1994
 A;Title: Antibodies generated from human immunoglobulin miniloci in transgenic mice.
 A;Reference number: S43956; MUID:94248036; PMID:8190629
 A;Accession: S43959
 A;Molecule type: DNA
 A;Residues: 1-4 <WAG>
 C;Keywords: immunoglobulin

Query Match 16.7%; Score 1; DB 2; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 Db 4 G 4

RESULT 61

A53284
 T-cell receptor beta 2 chain D region, Dbeta2 - rabbit
 C;Species: Oryctolagus cuniculus (domestic rabbit)
 C;Date: 02-May-1994 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
 C;Accession: A53284
 R;Harindranath, N.; Alexander, C.B.; Mage, R.G.
 Mol. Immunol. 28, 881-888, 1991
 A;Title: Evolutionarily conserved organization and sequences of germline diversity and
 A;Reference number: A53284; MUID:91342695; PMID:1678859
 A;Accession: A53284
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-4 <HAR>

A;Cross-references: GB:S60737; NID:g233916; PIDN:AAB19517.1; PID:g233917
 A;Note: sequence extracted from NCBI backbone (NCBIN:60737, NCBIP:60739)
 C;Keywords: T-cell receptor

Query Match 16.7%; Score 1; DB 2; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 Db 1 G 1

RESULT 62

B53284
 T-cell receptor beta 2 chain D region, Dbeta2 - rabbit
 C;Species: Oryctolagus cuniculus (domestic rabbit)
 C;Date: 02-May-1994 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
 C;Accession: B53284
 R;Harindranath, N.; Alexander, C.B.; Mage, R.G.
 Mol. Immunol. 28, 881-888, 1991
 A;Title: Evolutionarily conserved organization and sequences of germline diversity and

A;Reference number: A53284; MUID:91342695; PMID:1678859

A;Accession: B53284
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-4 <HAR>
 A;Cross-references: GB:S60737; NID:g233916; PIDN:AAB19518.1; PID:g233918
 A;Note: sequence extracted from NCBI backbone (NCBIN:60737, NCBIP:60738)
 C;Keywords: T-cell receptor

Query Match 16.7%; Score 1; DB 2; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 Db 3 G 3

RESULT 63

PT0696
 T-cell receptor beta chain V-D-J region (100-2N) - mouse (fragment)
 C;Species: Mus musculus (house mouse)
 C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 C;Accession: PT0635; PT0612; PT0545; PT0692; PT0552; PT0696
 R;Reeney, A.J.
 J. Exp. Med. 174, 115-124, 1991
 A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions
 A;Reference number: PT0509; MUID:91277601; PMID:1711558
 A;Accession: PT0635
 A;Status: translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-4 <FEZ>
 A;Experimental source: newborn thymus, strain BALB/c, 100-2N
 A;Accession: PT0612
 A;Status: translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-4 <FEZ>
 A;Experimental source: newborn thymus, strain BALB/c, 111-1J
 A;Accession: PT0545
 A;Status: translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-4 <FE3>
 A;Experimental source: day 18 fetal thymus, strain BALB/c, 126-1AD
 A;Accession: PT0692
 A;Status: translation not shown
 A;Molecule type: DNA
 A;Residues: 1-4 <FE4>
 A;Experimental source: day 18 fetal thymus, strain BALB/c, 140-1M
 A;Accession: PT0552
 A;Status: translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-4 <FE5>
 A;Experimental source: day 18 fetal thymus, strain BALB/c, 126-1CI
 A;Accession: PT0696
 A;Status: translation not shown
 A;Molecule type: DNA
 A;Residues: 1-4 <FE6>
 A;Experimental source: newborn thymus, strain BALB/c, 135-1AA
 C;Keywords: T-cell receptor

Query Match 16.7%; Score 1; DB 2; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
 Db 2 S 2

RESULT 64

PT0645
 T-cell receptor beta chain V-D-J region (111-1AI) - mouse (fragment)
 C;Species: Mus musculus (house mouse)

C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 A;Accession: PT0645; PT0626
 A;Status: translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-4 <FE1>
 J. Exp. Med. 174, 115-124, 1991
 A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
 C;Species: Mus musculus (house mouse)
 C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 A;Reference number: PT0509; MUID:91277601; PMID:1711558
 A;Accession: PT0645
 A;Status: translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-4 <FE2>
 A;Experimental source: newborn thymus, strain BALB/c, 120-2CJ
 C;Keywords: T-cell receptor

Query Match 16.7%; Score 1; DB 2; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6
 |
 Db 2 S 2

RESULT 65

PT0633
 T-cell receptor beta chain V-D-J region (120-2C) - mouse (fragment)
 C;Species: Mus musculus (house mouse)
 C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 A;Accession: PT0633
 A;Status: translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-4 <FE2>
 A;Experimental source: newborn thymus, strain BALB/c
 C;Keywords: T-cell receptor

Query Match 16.7%; Score 1; DB 2; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6
 |
 Db 2 S 2

RESULT 66

PT0711
 T-cell receptor beta chain V-D-J region (120-2J) - mouse (fragment)
 C;Species: Mus musculus (house mouse)
 C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 A;Accession: PT0607; PT0674; PT0678; PT0570; PT0711; PT0710
 R;Feeney, A.J.
 J. Exp. Med. 174, 115-124, 1991
 A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
 C;Species: Mus musculus (house mouse)
 C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 A;Reference number: PT0509; MUID:91277601; PMID:1711558
 A;Accession: PT0607
 A;Status: translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-4 <FE1>
 A;Experimental source: newborn thymus, strain BALB/c, 120-2J
 A;Accession: PT0674
 A;Status: translation not shown
 A;Molecule type: DNA
 A;Residues: 1-4 <FE2>
 A;Experimental source: day 18 fetal thymus, strain BALB/c, 140-1G

Query Match 16.7%; Score 1; DB 2; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6
 |
 Db 2 S 2

A;Accession: PT0678
 A;Status: translation not shown
 A;Molecule type: DNA
 A;Residues: 1-4 <FE3>
 A;Experimental source: day 18 fetal thymus, strain BALB/c, 154-1L
 A;Accession: PT0570
 A;Status: translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-4 <FE4>
 A;Experimental source: day 19 fetal thymus, strain BALB/c, 141-1I
 A;Accession: PT0711
 A;Status: translation not shown
 A;Molecule type: DNA
 A;Residues: 1-4 <FE5>
 A;Experimental source: newborn thymus, strain BALB/c (clones 161-2AE and 161-2AP)
 C;Keywords: T-cell receptor

Query Match 16.7%; Score 1; DB 2; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6
 |
 Db 2 S 2

RESULT 67

PT0661
 T-cell receptor beta chain V-D-J region (121-1BV) - mouse (fragment)
 C;Species: Mus musculus (house mouse)
 C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 A;Accession: PT0661
 A;Status: translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-4 <FE6>
 A;Experimental source: day 4 postnatal thymus, strain BALB/c
 C;Keywords: T-cell receptor

Query Match 16.7%; Score 1; DB 2; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6
 |
 Db 2 S 2

RESULT 68

PT0712
 T-cell receptor beta chain V-D-J region (121-3BC) - mouse (fragment)
 C;Species: Mus musculus (house mouse)
 C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 A;Accession: PT0649; PT0712
 R;Feeney, A.J.
 J. Exp. Med. 174, 115-124, 1991
 A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
 A;Reference number: PT0509; MUID:91277601; PMID:1711558
 A;Accession: PT0649
 A;Status: translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-4 <FE7>
 A;Experimental source: day 4 postnatal thymus, strain BALB/c, 121-3BC
 A;Accession: PT0712
 A;Status: translation not shown
 A;Molecule type: DNA
 A;Residues: 1-4 <FE8>
 A;Experimental source: newborn thymus, strain BALB/c, 165-3B
 C;Keywords: T-cell receptor

Query Match 16.7%; Score 1; DB 2; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6
 |
 Db 2 S 2

```

Query Match      16.7%; Score 1; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 H 2
      |
Db      4 H 4

RESULT 69
T-cell receptor beta chain V-D-J region (126-1AC) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0534
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0534
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-4 <FEE>
A:Experimental source: adult thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match      16.7%; Score 1; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      6 S 6
      |
Db      2 S 2

RESULT 70
T-cell receptor beta chain V-D-J region (126-1AH) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 05-Sep-1997
C:Accession: PT0546; PT0698; PT0583
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0546
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-4 <FEE>
A:Experimental source: day 18 fetal thymus, strain BALB/c, 126-1AH
A:Accession: PT0698
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-4 <FE2>
A:Experimental source: newborn thymus, strain BALB/c, 135-1BP
A:Accession: PT0583
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-4 <FE3>
A:Experimental source: day 19 fetal thymus, strain BALB/c, clone 159-2C
C:Keywords: T-cell receptor

Query Match      16.7%; Score 1; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      6 S 6
      |
Db      2 S 2

RESULT 71

```

```

PT0677
T-cell receptor beta chain V-D-J region (126-1BG) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0558; PT0677
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0558
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-4 <FEE>
A:Experimental source: day 18 fetal thymus, strain BALB/c, 126-1BG
A:Accession: PT0677
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-4 <FE2>
A:Experimental source: day 18 fetal thymus, strain BALB/c, 140-1BH
C:Keywords: T-cell receptor

Query Match      16.7%; Score 1; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      6 S 6
      |
Db      2 S 2

RESULT 72
PT0551
T-cell receptor beta chain V-D-J region (126-1CG) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0551
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0551
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-4 <FEE>
A:Experimental source: day 18 fetal thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match      16.7%; Score 1; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      6 S 6
      |
Db      2 S 2

RESULT 73
PT0697
T-cell receptor beta chain V-D-J region (135-1BF) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0697
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0697
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-4 <FEE>
A:Experimental source: newborn thymus, strain BALB/c
C:Keywords: T-cell receptor

```

Query Match 16.7%; Score 1; DB 2; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
 |
 Db 2 S 2

RESULT 74
 PT0706
 T-cell receptor beta chain V-D-J region (135-1DU) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 C:Accession: PT0706
 R:Feeney, A.J.
 J. Exp. Med. 174, 115-124, 1991
 A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
 A:Reference number: PT0509; PMID:1711558
 A:Accession: PT0706
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-4 <FEE>
 A:Experimental source: newborn thymus, strain BALB/c
 C:Keywords: T-cell receptor

Query Match 16.7%; Score 1; DB 2; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
 |
 Db 2 S 2

RESULT 75
 PT0675
 T-cell receptor beta chain V-D-J region (140-1AC) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 C:Accession: PT0675
 R:Feeney, A.J.
 J. Exp. Med. 174, 115-124, 1991
 A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
 A:Reference number: PT0509; PMID:1711558
 A:Accession: PT0675
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-4 <FEE>
 A:Experimental source: day 18 fetal thymus, strain BALB/c
 C:Keywords: T-cell receptor

Query Match 16.7%; Score 1; DB 2; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
 |
 Db 2 S 2

RESULT 76
 PT0721
 T-cell receptor beta chain V-D-J region (140-2K) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 C:Accession: PT0721
 R:Feeney, A.J.
 J. Exp. Med. 174, 115-124, 1991
 A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
 A:Reference number: PT0509; PMID:1711558
 A:Accession: PT0721
 A:Status: translation not shown

A:Molecule type: DNA
 A:Residues: 1-4 <FEE>
 A:Experimental source: newborn thymus, strain BALB/c
 C:Keywords: T-cell receptor

Query Match 16.7%; Score 1; DB 2; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
 |
 Db 2 S 2

RESULT 77
 PT0566
 T-cell receptor beta chain V-D-J region (141-1CH) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 C:Accession: PT0566
 R:Feeney, A.J.
 J. Exp. Med. 174, 115-124, 1991
 A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
 A:Reference number: PT0509; PMID:1711558
 A:Accession: PT0566
 A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-4 <FEE>
 A:Experimental source: day 19 fetal thymus, strain BALB/c
 C:Keywords: T-cell receptor

Query Match 16.7%; Score 1; DB 2; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
 |
 Db 2 S 2

RESULT 78
 A40135
 branched-chain-amino-acid transaminase (EC 2.6.1.42), mitochondrial - rat (fragment)
 N:Alternate names: branched-chain-amino-acid aminotransferase, mitochondrial
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 13-May-1992 #sequence_revision 13-May-1992 #text_change 30-Sep-1993
 C:Accession: A40135
 R:Hutson, S.M.; Wallin, R.; Hall, T.R.
 submitted to the Protein Sequence Database, March 1992
 A:Reference number: A40135
 A:Accession: A40135
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-4 <HUT>
 C:Keywords: aminotransferase; mitochondrion

Query Match 16.7%; Score 1; DB 2; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
 |
 Db 2 S 2

RESULT 79
 S47552
 ubiquitin - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 17-Mar-1999
 C:Accession: S47552
 R:Hubbard, M.J.; Carne, A.
 Biochim. Biophys. Acta 1200, 191-196, 1994

A:Title: Differential feeding-related regulation of ubiquitin and calbindin(9kDa) in rat
 A:Reference number: S47552; MUID:94304928; PMID:8031840
 A:Accession: S47552
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-4 <HUB>

Query Match 16.7%; Score 1; DB 2; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 |
 DB 3 G 3

RESULT 80

S09478
 Globulin IV alpha subunit gamma chain, seed - cucurbit (fragment)
 N:Alternate names: IIS globulin alpha subunit gamma chain
 C:Species: Cucurbita sp. (cucurbit)
 C:Date: 21-Nov-1993 #sequence_revision 08-Nov-1996 #text_change 08-Nov-1996
 A:Accession: S09478
 R:Omimiya, M.; Hara, I.; Matsubara, H.
 Plant Cell Physiol. 21, 157-167, 1980
 A:Title: Pumpkin (Cucurbita sp.) seed globulin IV. Terminal sequences of the acidic and
 A:Reference number: S09066
 A:Accession: S09478
 A:Molecule type: protein
 A:Residues: 1-4 <OHM>

Query Match 16.7%; Score 1; DB 2; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 |
 DB 3 G 3

RESULT 81

A60521
 Glycogen phosphorylase (EC 2.4.1.1), muscle - mullet (Liza ramada) (fragment)
 N:Alternate names: glycogen phosphorylase b
 C:Species: Liza ramada
 C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 11-May-2000
 A:Accession: A60521
 R:Bonamusa, L.; Baanante, I.V.
 Comp. Biochem. Physiol. B 95, 295-301, 1990
 A:Title: Purification and characterization of glycogen phosphorylase B from skeletal muscle
 A:Reference number: A60521; MUID:90227907; PMID:2109669
 A:Accession: A60521
 A:Molecule type: protein
 A:Residues: 1-5 <BON>

C:Superfamily: phosphorylase
 C:Keywords: glycosyltransferase; hexosyltransferase; phosphoprotein
 F:3/Binding site: phosphate (Ser) (covalent) (by phosphorylase b kinase) #status experimental

Query Match 16.7%; Score 1; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
 |
 DB 3 S 3

RESULT 82

JN0860
 peptidyl-dipeptidase A inhibitory peptide C107 - striped bonito
 C:Species: Sarda orientalis (striped bonito)
 C:Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 07-May-1999
 A:Accession: JN0860

R;Matsumura, N.; Fujii, M.; Takeda, Y.; Shimizu, T.
 Biosci. Biotechnol. Biochem. 57, 1743-1744, 1993
 A:Title: Isolation and characterization of angiotensin I-converting enzyme inhibitory
 A:Reference number: JN0859; MUID:94080036; PMID:7764272
 A:Accession: JN0860
 A:Molecule type: protein
 A:Residues: 1-5 <MAT>

A:Experimental source: intestine
 C:Comment: The carboxyl-terminus is essential for the protein's expression of angiotensin
 C:Superfamily: bradykinin-potentiating peptide
 C:Keywords: angiotensin-converting enzyme inhibitor

Query Match 16.7%; Score 1; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 H 2
 |
 DB 4 H 4

RESULT 83

A32516
 cholecystokinin-5 - dog
 N:Alternate names: CCK-5
 C:Species: Canis lupus familiaris (dog)
 C:Date: 18-Oct-1989 #sequence_revision 18-Oct-1989 #text_change 18-Aug-2000
 A:Accession: A32516
 R;Shively, J.; Reeve Jr., J.R.; Byssselein, V.E.; Ben-Avram, C.; Vigna, S.R.; Walsh, J.
 Am. J. Physiol. 252, G272-G275, 1987
 A:Title: CCK-5: sequence analysis of a small cholecystokinin from canine brain and in
 A:Reference number: A32516; MUID:87153871; PMID:3826354
 A:Accession: A32516
 A:Molecule type: protein
 A:Residues: 1-5 <SHI>

C:Comment: This peptide corresponds to the five carboxyl-terminal residues of cholecystokinin
 C:Superfamily: gastrin
 C:Keywords: amidated carboxyl end; neuropeptide
 F:5/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 16.7%; Score 1; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 |
 DB 1 G 1

RESULT 84

C23751
 spinal cord peptide SCP-6 - pig
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 28-Sep-1987 #sequence_revision 28-Sep-1987 #text_change 18-Aug-2000
 A:Accession: C23751
 R;Hsi, K.L.; Chen, R.L.; Chen, Z.G.; Zhang, H.L.; Lu, Y.A.; Guo, S.Y.; Wu, S.X.; Tsou
 Arch. Biochem. Biophys. 240, 178-183, 1985
 A:Reference number: A23751; MUID:85250425; PMID:4015098
 A:Accession: C23751
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-5 <HSI>

C:Superfamily: unassigned animal peptides

Query Match 16.7%; Score 1; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 |
 DB 3 G 3

RESULT 85

S70154
URF2 protein - Xanthomonas sp.
C:Species: Xanthomonas sp.
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 08-Oct-1999
C:Accession: S70154
R:Kholodii, G.Y.; Mindlin, S.Z.; Bass, I.A.; Yurieva, O.V.; Xinkhina, S.V.; Nikiforov, M.O. Microbiol. 17, 1189-1200, 1995
A:Title: Four genes, two ends, and a res region are involved in transposition of Tn5053:
A:Reference number: S70140; MUID:96130850; PMID:8594337
A:Accession: S70154
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-5 <RFO>
A:Cross-references: EMBL:140585; NID:G710572; PIDN:AAA98329.1; PID:G735909
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1995

Query Match 16.7%; Score 1; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
|
Db 4 S 4

RESULT 86

C41225
copper resistance protein - Pseudomonas syringae pv. tomato (fragment)
C:Species: Pseudomonas syringae pv. tomato
C:Date: 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change 24-Jun-1993
C:Accession: C41225
R:Cha, J.S.; Cooksey, D.A.
Proc. Natl. Acad. Sci. U.S.A. 88, 8915-8919, 1991
A:Title: Copper resistance in Pseudomonas syringae mediated by periplasmic and outer mem
A:Reference number: A41225; MUID:92020961; PMID:1924351
A:Accession: C41225
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-5 <CHA>

Query Match 16.7%; Score 1; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 H 2
|
b 1 H 1

RESULT 87

C41225
copper resistance protein A - Pseudomonas syringae pv. tomato (fragment)
C:Species: Pseudomonas syringae pv. tomato
C:Date: 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change 24-Jun-1993
C:Accession: A41225
R:Cha, J.S.; Cooksey, D.A.
Proc. Natl. Acad. Sci. U.S.A. 88, 8915-8919, 1991
A:Title: Copper resistance in Pseudomonas syringae mediated by periplasmic and outer mem
A:Reference number: A41225; MUID:92020961; PMID:1924351
A:Accession: A41225
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-5 <CHA>

Query Match 16.7%; Score 1; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
|
Db 2 S 2

RESULT 88

S70154
URF2 protein - Xanthomonas sp.
C:Species: Xanthomonas sp.
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 08-Oct-1999
C:Accession: S70154
R:Kholodii, G.Y.; Mindlin, S.Z.; Bass, I.A.; Yurieva, O.V.; Xinkhina, S.V.; Nikiforov, M.O. Microbiol. 17, 1189-1200, 1995
A:Title: Four genes, two ends, and a res region are involved in transposition of Tn5053:
A:Reference number: S70140; MUID:96130850; PMID:8594337
A:Accession: S70154
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-5 <RFO>
A:Cross-references: EMBL:140585; NID:G710572; PIDN:AAA98329.1; PID:G735909
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1995

Query Match 16.7%; Score 1; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 H 2
|
Db 5 H 5

RESULT 89

I40702
primase - Citrobacter diversus (fragment)
C:Species: Citrobacter diversus
C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 16-Aug-1996
C:Accession: I40702
R:Versalovic, J.; Lupski, J.R.
Mol. Microbiol. 8, 343-355, 1993
A:Title: Conservation and evolution of the rpsU-dnaG-rpoD macromolecular synthesis (WMS
A:Reference number: I40702; MUID:93302510; PMID:8316085
A:Accession: I40702
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-5 <RES>
A:Cross-references: GB:L01754; NID:gl44439
C:Genetics:
A:Gene: dnaG

Query Match 16.7%; Score 1; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
|
Db 3 G 3

RESULT 90

B37325
pap fibrial regulatory protein papI - Escherichia coli (fragment)
C:Species: Escherichia coli
C:Date: 11-Sep-1992 #sequence_revision 11-Sep-1992 #text_change 23-Mar-1993
C:Accession: B37325
R:Bratzen, B.A.; Blyn, L.B.; Skinner, B.S.; Low, D.A.
J. Bacteriol. 173, 1789-1800, 1991
A:Title: Evidence for a methylation-blocking factor (mbf) locus involved in pap pilus ex
A:Reference number: A37325; MUID:91154136; PMID:1671857
A:Accession: B37325
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-5 <BRA>
A:Cross-references: GB:M63747

Query Match 16.7%; Score 1; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      6 S 6
      1
Db      2 S 2

RESULT 91
A44955
alkanal monooxygenase (FMN-linked) (EC 1.14.14.3) alpha chain - Vibrio harveyi (fragment)
C:Species: Vibrio harveyi
C>Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 26-May-2000
C:Accession: A44955
R:Paquatte, O.; Tu, S. C.
Photochem. Photobiol. 50, 817-825, 1989
A:Title: Chemical modification and characterization of the alpha cysteine 106 at the Vib
A:Reference number: A44955; MUID:90175700; PMID:2626493
A:Accession: A44955
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-5 <PAQ>
C:Keywords: FMN; luminescence; monooxygenase; oxidoreductase

Query Match      16.7%; Score 1; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 G 1
      1
Db      2 G 2

RESULT 92
B31836
20K protein - Rickettsia rickettsii (fragment)
C:Species: Rickettsia rickettsii
C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 28-May-1999
C:Accession: B31836
R:Anderson, B.E.; Baumstark, B.R.; Bellini, W.J.
J. Bacteriol. 170, 4493-4500, 1988
A:Title: Expression of the gene encoding the 17-kilodalton antigen from Rickettsia rick
A:Reference number: A91895; MUID:89008059; PMID:3119629
A:Accession: B31836
A:Molecule type: DNA
A:Residues: 1-5 <AND>
A:Cross-references: GB:J03371; NID:G152455; PIDN:AAD15030.1; PID:G4262874

Query Match      16.7%; Score 1; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 G 1
      1
Db      2 G 2

RESULT 93
I40469
dnazX-like protein - Bacillus subtilis (fragment)
C:Species: Bacillus subtilis
C>Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
C:Accession: I40469
R:Struck, J.C.; Hartmann, R.K.; Toschka, H.Y.; Erdmann, V.A.
Mol. Gen. Genet. 215, 478-482, 1989
A:Title: Transcription and processing of Bacillus subtilis small cytoplasmic RNA.
A:Reference number: I40469; MUID:89218958; PMID:2468993
A:Accession: I40469
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-5 <RES>
A:Cross-references: EMBL:X14796; NID:G40130; PIDN:CAA32902.1; PID:G4376204
C:Genetics:
A:Start codon: GTG

Query Match      16.7%; Score 1; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      6 S 6
      1
Db      5 S 5

RESULT 94
D60274
major protein antigen MPT46 - Mycobacterium tuberculosis (fragment)
C:Species: Mycobacterium tuberculosis
C>Date: 11-Dec-1992 #sequence_revision 11-Dec-1992 #text_change 30-Sep-1993
C:Accession: D60274
R:Nagai, S.; Wiker, H.G.; Harboe, M.; Kinomoto, M.
Infect. Immun. 59, 372-382, 1991
A:Title: Isolation and partial characterization of major protein antigens in the cult
A:Reference number: A60274; MUID:9109989; PMID:1898899
A:Accession: D60274
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-5 <NAG>

Query Match      16.7%; Score 1; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      6 S 6
      1
Db      3 S 3

RESULT 95
F22565
R-phycocerythrin gamma-A chain - red alga (Gastrocionium coulteri) (fragment)
C:Species: Gastrocionium coulteri
C>Date: 07-Mar-1988 #sequence_revision 07-Mar-1988 #text_change 23-Mar-1993
C:Accession: F22565
R:Klotz, A.V.; Glazer, A.N.
J. Biol. Chem. 260, 4856-4863, 1985
A:Title: Characterization of the bilin attachment sites in R-phycocerythrin.
A:Reference number: A22565; MUID:85182601; PMID:3886644
A:Accession: F22565
A:Molecule type: protein
A:Residues: 1-5 <KLO>

Query Match      16.7%; Score 1; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 G 1
      1
Db      1 G 1

RESULT 96
T14908
hypothetical protein - parsley
C:Species: Petroselinum crispum (parsley)
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T14908
R:Kircher, S.; Ledger, S.; Hayaishi, H.; Weisshaar, B.; Schafer, E.; Frohnmeyer, H.
Mol. Gen. Genet. 257, 595-605, 1998
A:Title: CPRF4a, a novel plant bZIP protein of the CPRF family: comparative analysis c
A:Reference number: Z18261; MUID:98265918; PMID:9604882
A:Accession: T14908
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-5 <KIR>
A:Cross-references: EMBL:Y10809; NID:G3336901; PIDN:CAA71767.1; PID:G3336902
A:Experimental source: Hamburger Schlutt

```

Query Match 16.7%; Score 1; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
 |
 Db 3 S 3

RESULT 97

T14910
 hypothetical protein - parsley
 C:Species: Petroselinum crispum (parsley)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
 C:Accession: T14910
 R:Kircher, S.; Ledger, S.; Hayashi, H.; Weisshaar, B.; Schafer, E.; Frohnmeyer, H.
 Mol. Gen. Genet. 257, 595-605, 1998
 A:Title: CPRP4a, a novel plant bZIP protein of the CPRF family: comparative analysis of
 A:Reference number: Z18261; MUID:98265918; PMID:9604882
 A:Accession: T14910
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-5 <KIP>
 A:Cross-references: EMBL:Y10810; NID:g3336904; PIDN:CAA1769.1; PID:g3336905
 A:Experimental source: ssp. Hamburger Schnitt

Query Match 16.7%; Score 1; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
 |
 Db 3 S 3

RESULT 98

A33882
 cadmium-binding pentapeptide - downy thornapple
 C:Species: Datura innoxia (downy thornapple)
 C:Date: 21-May-1990 #sequence_revision 21-May-1990 #text_change 18-Jun-1993
 C:Accession: A33882
 R:Jackson, P.J.; Unkefer, C.J.; Doolen, J.A.; Watt, K.; Robinson, N.J.
 Proc. Natl. Acad. Sci. U.S.A. 84, 6619-6623, 1987
 A:Title: Poly(gamma-glutamylcysteinyl)glycine: its role in cadmium resistance in plant
 A:Reference number: A94182; MUID:88016144; PMID:3477793
 A:Accession: A33882
 A:Molecule type: protein
 A:Residues: 1-5 <JAC>

Query Match 16.7%; Score 1; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 |
 Db 5 G 5

RESULT 99

PQ0689
 photosystem I 10.4K H1 chain - common tobacco (fragment)
 C:Species: Nicotiana tabacum (common tobacco)
 C:Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 17-Mar-1999
 C:Accession: PQ0689
 R:Obokata, J.; Mikami, K.; Hayaehida, N.; Nakamura, M.; Sugiyura, M.
 Plant Physiol. 102, 1259-1267, 1993
 A:Title: Molecular heterogeneity of photosystem I. psaD, psaE, psaF, psaH and psaL are a
 A:Reference number: PQ0667; MUID:94105345; PMID:8278548
 A:Accession: PQ0689
 A:Molecule type: protein
 A:Residues: 1-5 <OBO>
 C:Keywords: chloroplast; photosynthesis; photosystem I; thylakoid

Query Match 16.7%; Score 1; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 |
 Db 3 G 3

RESULT 100

S51077
 alpha-amylase - rice
 C:Species: Oryza sativa (rice)
 C:Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 01-Sep-1995
 C:Accession: S51077
 R:Terashima, M.; Kubo, A.; Suzawa, M.; Itoh, Y.; Katoh, S.
 Eur. J. Biochem. 226, 249-254, 1994
 A:Title: The roles of the N-linked carbohydrate chain of rice alpha-amylase in thermosta
 A:Reference number: S51077; MUID:95045597; PMID:7957256
 A:Accession: S51077
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-5 <YER>

Query Match 16.7%; Score 1; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
 |
 Db 1 S 1

Search completed: October 21, 2003, 18:29:08

Job time : 21.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:12:30 ; Search time 10.3333 Seconds

(without alignments)
27.306 Million cell updates/sec

Title: US-10-057-890A-15

Perfect score: 6

Sequence: 1 GHHHS 6

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 372

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Listing first 150 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2	33.3	3	1 GRWM HUMAN	P01157 homo sapien
2	2	33.3	4	1 DCM_L_PSECH	P15916 pseudomonas
3	2	33.3	5	1 B10B CITRUS	P12997 citrobacter
4	2	33.3	7	1 CHOX ALCSP	P16101 alcaligenes
5	2	33.3	7	1 HY7 PIG	P01153 sus scrofa
6	2	33.3	8	1 GLVR HUMAN	P02729 homo sapien
7	2	33.3	8	1 LCK4 LEUMA	P21143 leucophaea
8	2	33.3	8	1 LCK6 LEUMA	P19988 leucophaea
9	2	33.3	8	1 UF06 MOUSE	P38644 mus musculus
10	2	33.3	9	1 BS43 SERPL	P83375 serratia pl
11	2	33.3	9	1 LITO LITAU	P08945 litoria aur
12	2	33.3	9	1 LITR PHYRO	P08946 phylomedus
13	2	33.3	9	1 NSK1 SARBU	P41492 sarcophaga
14	2	33.3	10	1 FARP WANSE	P18523 manduca sex
15	2	33.3	10	1 GLEM HUMAN	P02728 homo sapien
16	2	33.3	10	1 GRP_RANRI	P23260 rana ridibu
17	2	33.3	10	1 LSK2 LEUMA	P09039 leucophaea
18	1	16.7	3	1 THYL PTG	P01151 sus scrofa
19	1	16.7	4	1 ACHI ACHFU	P35904 achatina fu
20	1	16.7	4	1 EOST HUMAN	P02731 homo sapien
21	1	16.7	4	1 OCPI OCTMI	P58648 octopus min
22	1	16.7	4	1 OCPI OCTMI	P58649 octopus min
23	1	16.7	4	1 RM01 YEAST	P36515 saccharomyc
24	1	16.7	5	1 ALI4 CARMA	P81817 carcinus ma
25	1	16.7	5	1 E103 LITRU	P82099 litoria rub
26	1	16.7	5	1 E104 LITRU	P82100 litoria rub
27	1	16.7	5	1 PAP2 PARMA	P81864 pardachirus
28	1	16.7	5	1 RE32 LITRU	P82073 litoria rub
29	1	16.7	5	1 TPIS CANFA	P54714 canis famli
30	1	16.7	5	1 UF01 MOUSE	P38639 mus musculus
31	1	16.7	5	1 UX44 CHLTR	P38005 chlamydia t
32	1	16.7	6	1 CIP1 MYTID	P13736 mytilus edu
33	1	16.7	6	1 CIP2 MYTID	P13737 mytilus edu

34	1	16.7	6	1 FARP MONEX	P41966 moniezia ex
35	1	16.7	6	1 LOKI LOEMI	P41491 locusta mig
36	1	16.7	6	1 TMOF SARBU	P41495 sarcophaga
37	1	16.7	7	1 TRPI PSEPU	P36414 pseudomonas
38	1	16.7	7	1 ALI2 CARMA	P81805 carcinus ma
39	1	16.7	7	1 ALI3 CARMA	P81806 carcinus ma
40	1	16.7	7	1 ALI4 CARMA	P81807 carcinus ma
41	1	16.7	7	1 ALI5 CARMA	P81808 carcinus ma
42	1	16.7	7	1 ALI7 CYDPO	P82158 cydia pomon
43	1	16.7	7	1 BRHP CONIM	P58803 conus imper
44	1	16.7	7	1 CIA ENTFA	P11932 enterococcu
45	1	16.7	7	1 FARI HELTI	P41871 helisoma tr
46	1	16.7	7	1 FARP ASCSU	P31890 ascaris suu
47	1	16.7	7	1 FARP HAECO	P81298 haemonchus
48	1	16.7	7	1 FARP PANRE	P41874 panagrellus
49	1	16.7	7	1 FARP HIRME	P42564 hirudo medi
50	1	16.7	7	1 GFRP MOUSE	P99025 mus musculu
51	1	16.7	7	1 IGAO DACDE	P06294 dactylium d
52	1	16.7	7	1 LANC CARUI	P36960 carnobacter
53	1	16.7	7	1 MNPI LEPDE	P42984 leptinotars
54	1	16.7	7	1 PP2 LYCES	P83379 lycopersico
55	1	16.7	7	1 TY51 LITRU	P82065 litoria rub
56	1	16.7	7	1 UC24 MAIZE	P80630 zea mays fm
57	1	16.7	7	1 UF03 MOUSE	P38641 mus musculu
58	1	16.7	7	1 UF04 MOUSE	P38642 mus musculu
59	1	16.7	7	1 UH11 RAT	P56576 rattus norv
60	1	16.7	7	1 UN06 PINPS	P81675 pinus pinas
61	1	16.7	7	1 WWA1 ACHFU	P35919 achatina fu
62	1	16.7	7	1 WWA2 ACHFU	P35920 achatina fu
63	1	16.7	7	1 WWA3 ACHFU	P35921 achatina fu
64	1	16.7	8	1 ACI THUAL	P18691 thunnus alb
65	1	16.7	8	1 ACHG GRYBI	P14086 gryllus bim
66	1	16.7	8	1 AKH LIBAU	P25418 libellula a
67	1	16.7	8	1 AKH MELWL	P25423 melolontha
68	1	16.7	8	1 AKH TABAT	P14595 tabanus atr
69	1	16.7	8	1 ALI2 CARMA	P81815 carcinus ma
70	1	16.7	8	1 ALI5 CARMA	P81818 carcinus ma
71	1	16.7	8	1 ALI6 CARMA	P81819 carcinus ma
72	1	16.7	8	1 ALI7 CARMA	P81820 carcinus ma
73	1	16.7	8	1 ALI8 CARMA	P81821 carcinus ma
74	1	16.7	8	1 ALI1 CYDPO	P82152 cydia pomon
75	1	16.7	8	1 ALI3 CYDPO	P82154 cydia pomon
76	1	16.7	8	1 ALI4 CALVO	P41840 calliphora
77	1	16.7	8	1 ALI4 CYDPO	P82155 cydia pomon
78	1	16.7	8	1 ALI5 CALVO	P41841 calliphora
79	1	16.7	8	1 ALI5 CYDPO	P82156 cydia pomon
80	1	16.7	8	1 ALI6 CYDPO	P82157 cydia pomon
81	1	16.7	8	1 ALI7 CARMA	P81809 carcinus ma
82	1	16.7	8	1 ALI8 CARMA	P81811 carcinus ma
83	1	16.7	8	1 ALI9 CARMA	P81812 carcinus ma
84	1	16.7	8	1 ANG2 BOTJA	Q10582 bothriops ja
85	1	16.7	8	1 CAD1 ENTFA	P13268 enterococcu
86	1	16.7	8	1 CCKN MACEU	P30169 macropus eu
87	1	16.7	8	1 CLP THICU	P80488 thiobacillu
88	1	16.7	8	1 COW2 CONPU	P58785 conus purpu
89	1	16.7	8	1 COXG RAT	P80430 rattus norv
90	1	16.7	8	1 CPDI ENTFA	P13269 enterococcu
91	1	16.7	8	1 FARI PANRE	P41872 panagrellus
92	1	16.7	8	1 FARP PENMO	P83316 penaeus mon
93	1	16.7	8	1 FARP HOMAM	P41486 homarus ame
94	1	16.7	8	1 FARP ASCSU	P43171 ascaris suu
95	1	16.7	8	1 FARP CALVO	P41863 calliphora
96	1	16.7	8	1 FUSS FUSSO	P81010 fusarium so
97	1	16.7	8	1 HTF1 PERAM	P04548 periplaneta
98	1	16.7	8	1 HTF TENMO	P25419 tenebrio mo
99	1	16.7	8	1 LCK1 LEUMA	P21140 leucophaea
100	1	16.7	8	1 LCK2 LEUMA	P21141 leucophaea
101	1	16.7	8	1 LCK3 LEUMA	P21142 leucophaea
102	1	16.7	8	1 LCK5 LEUMA	P19987 leucophaea
103	1	16.7	8	1 LCK7 LEUMA	P19989 leucophaea
104	1	16.7	8	1 LCK8 LEUMA	P19990 leucophaea
105	1	16.7	8	1 LMT2 LOEMI	P22396 locusta mig
106	1	16.7	8	1 LPK LEUMA	P13049 leucophaea

107 1 16.7 8 1 LPMS STREP
 108 1 16.7 8 1 ORMY_ORCLI
 109 1 16.7 8 1 PPK2_PERAM
 110 1 16.7 8 1 RPCH_PANBO
 111 1 16.7 8 1 R31_ERWCH
 112 1 16.7 8 1 RS7_MYCIT
 113 1 16.7 8 1 RT34_BOVIN
 114 1 16.7 8 1 UH09_RAT
 115 1 16.7 8 1 UPAA1_HUMAN
 116 1 16.7 8 1 VGLG_HSV2B
 117 1 16.7 8 1 WPI_PERAT
 118 1 16.7 8 1 AL10_CARMA
 119 1 16.7 8 1 AL11_CARMA
 120 1 16.7 8 1 ALC_GHRE
 121 1 16.7 8 1 CCAP_CARMA
 122 1 16.7 8 1 CONO_CONGE
 123 1 16.7 8 1 CONO_CONST
 124 1 16.7 8 1 COW_CONVE
 125 1 16.7 8 1 D1_NEPNO
 126 1 16.7 8 1 DNFI_LOCHI
 127 1 16.7 8 1 DSIP_RABIT
 128 1 16.7 8 1 FAR2_CALVO
 129 1 16.7 8 1 FAR2_PANRE
 130 1 16.7 8 1 FAR3_CALVO
 131 1 16.7 8 1 FAR3_PENNO
 132 1 16.7 8 1 FAR4_PENNO
 133 1 16.7 8 1 FAR5_ASCSU
 134 1 16.7 8 1 FAR5_CALVO
 135 1 16.7 8 1 FAR5_PENNO
 136 1 16.7 8 1 FAR6_CALVO
 137 1 16.7 8 1 FAR6_MACRS
 138 1 16.7 8 1 FAR7_CALVO
 139 1 16.7 8 1 FAR8_MACRS
 140 1 16.7 8 1 FAR9_ASCSU
 141 1 16.7 8 1 FAR9_CALVO
 142 1 16.7 8 1 FAR9_CALSI
 143 1 16.7 8 1 FIBB_ERYP
 144 1 16.7 8 1 FIBB_ERYP
 145 1 16.7 8 1 FIBB_PAPAN
 146 1 16.7 8 1 FIBB_PAPHA
 147 1 16.7 8 1 FIBB_THEGE
 148 1 16.7 8 1 FRFI_SARBU
 149 1 16.7 8 1 HUTU_KLEAE
 150 1 16.7 8 1

ALIGNMENTS

RESULT 1
 ID GRWM HUMAN STANDARD; PRT; 3 AA.
 AC P01157;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 21-JUL-1986 (Rel. 01, Last annotation update)
 DE Growth-modulating peptide.
 DS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 DX NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE=77162369; PubMed=858356;
 RX Schlessinger D.H., Pickart L., Thaler M.M.;
 RT "Growth-modulating serum tripeptide is glycyl-histidyl-lysine.";
 RL Experientia 33:324-325(1977).
 CC -!- MISCELLANEOUS: THIS SERUM TRIPEPTIDE HAS BEEN FOUND TO STIMULATE
 CC GROWTH OF SOME CELL TYPES AND TO INHIBIT OTHER TYPES IN VITRO.
 DR GO: 0001558; P: regulation of cell growth; NAS.
 SQ SEQUENCE 3 AA; 340 MW; 6331E81000000000 CRC64;

Query Match

33.3%; Score 2; DB 1; Length 3;

Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GH 2
 DB 1 GH 2

RESULT 2

DCML_PSECH STANDARD; PRT; 4 AA.
 ID DCML_PSECH
 AC P19916;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Carbon monoxide dehydrogenase large chain (EC 1.2.99.2) (CO
 CUTL).
 GN CUTL.
 OS Pseudomonas carboxydohydrogena.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Bradyrhizobiaceae.
 OX NCBI_TaxID=290;
 RN [1]
 RP MEDLINE=90055678; PubMed=2818128;
 RX Kraut M., Hugendieck I., Herwig S., Meyer O.;
 RT "Homology and distribution of CO dehydrogenase structural genes in
 RT carboxydohydrogenic bacteria";
 RL Arch. Microbiol. 152:335-341(1989).
 CC -!- FUNCTION: Catalyzes the oxidation of carbon monoxide to carbon
 CC dioxide.
 CC -!- CATALYTIC ACTIVITY: CO + H(2)O + acceptor = CO(2) + reduced
 CC acceptor.
 CC -!- COFACTOR: Molybdenum (molybdopterin).
 CC -!- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
 CC SMALL.
 CC PIR; PLO140; PLO140.
 DR Oxidoreductase; Molybdenum.
 FT NON TER 4 4
 SQ SEQUENCE 4 AA; 441 MW; 7761E876F000000000 CRC64;

Query Match 33.3%; Score 2; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GH 2
 DB 2 GH 3

RESULT 3

BIOB_CITFR STANDARD; PRT; 5 AA.
 AC P12937;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Biotin synthase (EC 2.8.1.6) (Biotin synthetase) (Fragment).
 GN BIOB.
 OS Citrobacter freundii.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Citrobacter.
 OX NCBI_TaxID=546;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89006280; PubMed=2971595;
 RA Shuan D., Campbell A.;
 RT "Transcriptional regulation and gene arrangement of Escherichia coli,
 RT Citrobacter freundii and Salmonella typhimurium biotin operons.";
 RL Gene 67:203-211(1988).
 CC -!- CATALYTIC ACTIVITY: Dethiobiotin + sulfur = biotin.
 CC -!- PATHWAY: Biotin biosynthesis; last step.
 CC -!- SIMILARITY: BELONGS TO THE BIOTIN AND LIPOIC ACID SYNTHETASES

```

CC      FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M21922; -; NOT_ANNOTATED_CDS.
DR      PIR; I40698; I40698.
KW      Biotin biosynthesis; Iron-sulfur; Transferase.
FT      NON_TER 5
SQ      SEQUENCE 5 AA; 532 MW; 75A5B1EDD6F00000 CRC64;

Query Match 33.3%; Score 2; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 HS 6
DB      3 HS 4

RESULT 4
CHOX_ALCSP STANDARD; PRT; 7 AA.
AC      P16101;
DT      01-APR-1990 (Rel. 14, Created)
DT      01-APR-1990 (Rel. 14, Last sequence update)
DT      01-APR-1990 (Rel. 14, Last annotation update)
DE      Choline oxidase (EC 1.1.3.17) (Fragment).
OS      Alcaligenes sp.
OC      Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC      Alcaligenaceae; Alcaligenes.
OX      NCBI_TaxID=512;
RN      [1]
RP      SEQUENCE.
RX      MEDLINE=81006769; PubMed=6997203;
RA      Ohta-Fukuyama M., Miyake Y., Emi S., Yamano T.;
RT      "Identification and properties of the prosthetic group of choline
RT      oxidase from Alcaligenes sp.;"
RL      J. Biochem. 88:197-203(1980).
CC      -|- CATALYTIC ACTIVITY: Choline + O(2) = betaine aldehyde + H(2)O(2).
DR      PIR; A15398; A15398.
KW      Oxidoreductase.
FT      NON_TER 7
SQ      SEQUENCE 7 AA; 839 MW; 7415B1E457644AC0 CRC64;

Query Match 33.3%; Score 2; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 HS 6
DB      5 HS 6

RESULT 5
HY7_PIG STANDARD; PRT; 7 AA.
AC      P01153;
DT      21-JUL-1986 (Rel. 01, Created)
DT      21-JUL-1986 (Rel. 01, Last sequence update)
DT      21-JUL-1986 (Rel. 01, Last annotation update)
DE      Hypothalamic heptapeptide.
OS      Sus scrofa (Pig).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Euthera; Cetartiodactyla; Suina; Suidae; Sus.
OX      NCBI_TaxID=9823;
RN      [1]
RP      SEQUENCE, AND SYNTHESIS.

```

```

RX      MEDLINE=81213980; PubMed=6263778;
RA      Chang R.C.C., Huang W.-Y., Arimura A., Redding T.W., Coy D.H.,
RA      Saffran M., Kong A., Hamilton J.W., Cohn D.V., Schally A.V.;
RT      "Isolation, structure and synthesis of a heptapeptide with in vitro
RT      ACTH-releasing activity from porcine hypothalamus.;"
RL      Horm. Metab. Res. 13:228-232(1981).
DR      PIR; A01417; NYPG7.
SQ      SEQUENCE 7 AA; 957 MW; 632B45B1FB5059A0 CRC64;

Query Match 33.3%; Score 2; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 HS 6
DB      4 HS 5

RESULT 6
GLUR_HUMAN STANDARD; PRT; 8 AA.
AC      P02729;
DT      21-JUL-1986 (Rel. 01, Created)
DT      21-JUL-1986 (Rel. 01, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Urine glycopeptide.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE.
RX      MEDLINE=72062338; PubMed=5126885;
RA      Lote C.J., Weiss J.B.;
RT      "Identification in urine of a low-molecular-weight highly polar
RT      glycopeptide containing cysteinyl-galactose.;"
RL      Biochem. J. 123:25P-25P(1971).
CC      -|- FUNCTION: THE IDENTITY OF THE GLYCOPROTEIN FROM WHICH THIS PEPTIDE
CC      IS DERIVED IS UNKNOWN. NO PHYSIOLOGICAL FUNCTION HAS BEEN
CC      ATTRIBUTED. AN ERYTHROCYTE MEMBRANE GLYCOPOLYPEPTIDE HAVING A
CC      SIMILAR STRUCTURE HAS ALSO BEEN FOUND.
DR      PIR; A03188; XGHUEU.
DR      GO; GO:0005576; C:extracellular; NAS.
KW      Glycoprotein.
FT      CARBOHYD 1 1 S-LINKED (GAL. .).
SQ      SEQUENCE 8 AA; 855 MW; C2D87AA1F5B1EB1E CRC64;

Query Match 33.3%; Score 2; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 HS 6
DB      3 HS 4

RESULT 7
LCK4_LEUMA STANDARD; PRT; 8 AA.
AC      P21143;
DT      01-MAY-1991 (Rel. 18, Created)
DT      01-MAY-1991 (Rel. 18, Last sequence update)
DT      01-MAY-1991 (Rel. 18, Last annotation update)
DE      Leucokinin IV (I-IV).
OS      Leucophaea maderae (Madeira cockroach).
OC      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC      Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;
OC      Blaberidae; Leucophaea.
OX      NCBI_TaxID=6988;
RN      [1]
RP      SEQUENCE, AND SYNTHESIS.
RC      TISSUE=Head;
RA      Holman G.M., Cook B.J., Nachman R.J.;

```

RT "Primary structure and synthesis of two additional neuropeptides
 RT from Leucophaea maderae: members of a new family of
 RT Cephalomyotropins.";
 RL Comp. Biochem. Physiol. 84C:271-276(1986).
 CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
 CC ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).
 CC -!- SIMILARITY: TO THE OTHER LEUCOKININS.
 CC Neuropeptide; Amidation.
 CC KW MOD RES 8 8 AMIDATION.
 CC FT SEQUENCE 8 AA; 906 MW; DC6365B1E9D58DDA CRC64;

Query Match 33.3%; Score 2; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 HS 6
 ||
 Db 5 HS 6

RESULT 8

CK6_LEUMA
 ID D LCK6_LEUMA STANDARD; PRT; 8 AA.
 AC P19988;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Leucokinin VI (L-VI).
 CS Leucophaea maderae (Madeira cockroach).
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;
 CC Blaberidae; Leucophaea.
 CC NCBI_TaxID=6988;
 CC [1]

RP SEQUENCE.

RT TISSUE=Head;
 RC MEDLINE=87052651; PubMed=2877794;
 RA Holman G.M., Cook B.J., Nachman R.J.;
 RA "Isolation, primary structure, and synthesis of leucokinin V and VI:
 RT myotropic peptides of Leucophaea maderae.";
 RL Comp. Biochem. Physiol. 88C:271-30(1987).
 CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
 CC ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).
 CC -!- SIMILARITY: TO THE OTHER LEUCOKININS, AND TO MANDUCA SEXTA AND
 CC HELIOTHIS ZEA ADIPOKINETIC HORMONE.
 CC PIR, JS0316; JS0316.
 CC Neuropeptide; Amidation; Pyrrolidone carboxylic acid.
 CC MOD RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 CC T MOD RES 8 8 AMIDATION.
 CC Q SEQUENCE 8 AA; 935 MW; 9DG365B1E9D5A5A6 CRC64;

Query Match 33.3%; Score 2; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 HS 6
 ||
 Db 5 HS 6

RESULT 9

JF06_MOUSE
 ID ID_UF06_MOUSE STANDARD; PRT; 8 AA.
 AC P38644;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE Unknown protein from 2D-page of fibroblasts (P50) (Fragment).
 CS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID=10090;
 CC [1]

RP SEQUENCE.
 RC TISSUE=Fibroblast;
 RX MEDLINE=95009907; PubMed=7523108;
 RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
 RA "Separation and sequencing of familial and novel murine proteins
 RT using preparative two-dimensional gel electrophoresis.";
 RL Electrophoresis 15:735-745(1994).
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 5.2, ITS MW IS: 50 kDa.
 CC NON TER 8 8
 CC FT SEQUENCE 8 AA; 817 MW; A35DD878676B05B1 CRC64;

Query Match 33.3%; Score 2; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 HS 6
 ||
 Db 1 HS 2

RESULT 10

BS43_SERPL
 ID BS43_SERPL STANDARD; PRT; 9 AA.
 AC P83375;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Bacteriocin serracin P 43 kDa subunit (Fragment).
 OS Serratia plymuthica.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Serratia.
 CC NCBI_TaxID=82996;
 CC [1]

RP SEQUENCE, AND FUNCTION.

RC STRAIN=J7;
 RX MEDLINE=2293561; PubMed=12406768;
 RA Jabrane A., Sabri A., Compere P., Jacques P., Vandenberghe I.,
 RA Van Beeumen J., Thonart P.;
 RA "Characterization of serracin P, a phage-tail-like bacteriocin, and
 RT its activity against *Erwinia amylovora*, the fire blight pathogen.";
 RL Appl. Environ. Microbiol. 68:5704-5710(2002).
 CC -!- FUNCTION: Major component of a prophage tail sheath (Probable).
 CC -!- FUNCTION: Antibacterial activity against Gram-negative bacterium
 CC *E. amylovora*.
 CC InterPro; IPR006498; Tail_tube.
 CC PIRam; PF04985; Phage_tube; 1.
 CC KW Antibiotic; Bacteriocin.
 CC FT NON TER 9 9
 CC Q SEQUENCE 9 AA; 1095 MW; 1E66D412C871E1FB CRC64;

Query Match 33.3%; Score 2; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HH 3
 ||
 Db 3 HH 4

RESULT 11

LITO_LITAU
 ID LITO_LITAU STANDARD; PRT; 9 AA.
 AC P08925;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Litorin.
 OS Litoria aurea (Green and golden bell frog).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
 CC Pelodyadinae; Litoria.
 CC NCBI_TaxID=8371;

CC -!- FUNCTION: INCREASES THE FORCE OF NEURALLY EVOKED CONTRACTIONS IN
 CC THE MAJOR POWER-PRODUCING FLIGHT MUSCLES, THE DORSAL LONGITUDINAL
 CC MUSCLES AND SO IS LIKELY TO PLAY A ROLE IN SUSTAINING OR PROMOTING
 CC FLIGHT BEHAVIOR PATTERNS.

CC -!- SIMILARITY: BELONGS TO THE FARP (FMRPAMIDE RELATED PEPTIDE)
 CC FAMILY.

CC PIR: A43977; A43977.
 CC Amidation; Neuropeptide; Pyrrolidone carboxylic acid.
 KW MOD RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD RES 10 10 AMIDATION.
 SQ SEQUENCE 10 AA; 1247 MW; D3C45229D5B1F2D2 CRC64;

Query Match 33.3%; Score 2; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 2e+03; 0; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 HS 6

DB 5 HS 6

RESULT 15

GLEM HUMAN

ID GLEM HUMAN STANDARD; PRT; 10 AA.

AC P02728;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Erythrocyte membrane glycopeptide.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE.

RX MEDLINE=72034940; PubMed=5286858;

RA Weiss J.B., Lote C.J., Bobinski H.;

RT "New low molecular weight glycopeptide containing triglycosylcysteine

RT in human erythrocyte membrane.";

RL Nature New Biol. 234:25-26(1971);

CC -!- PFM: S-LINKED GLYCAN CONSISTS OF GLC-GLC-GLC TRISACCHARIDE.

CC -!- MISCELLANEOUS: THE IDENTITY OF THE GLYCOPROTEIN FROM WHICH THIS

CC PEPTIDE IS DERIVED IS UNKNOWN. NO PHYSIOLOGICAL FUNCTION HAS BEEN

CC ATTRIBUTED.

CC PIR: A03187; XGHUE.

KW Glycoprotein; Erythrocyte.

FT CARBOHYD 1 1 S-LINKED (GLC...).

SQ SEQUENCE 10 AA; 1049 MW; 239BFEEA1F5B1E8 CRC64;

Query Match

Best Local Similarity 100.0%; Pred. No. 2e+03; 0; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GH 2

DB 3 GH 4

RESULT 16

GRP_RANRI

ID GRP_RANRI STANDARD; PRT; 10 AA.

AC P23260;

DT 01-NOV-1991 (Rel. 20, Created)

DT 01-NOV-1991 (Rel. 20, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Neuromedin C.

OS Rana ridibunda (Laughing frog) (Marsh frog).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Rana.

OX NCBI_TaxID=8406;

RN [1]

RP SEQUENCE.

CC TISSUE=Brain;

RX MEDLINE=91315477; PubMed=1859413;
 RA Conlon J.M., O'Harte F., Vaudry H.;

RT "Primary structures of the bombesin-like neuropeptides in frog brain

RT show that bombesin is not the amphibian gastrin-releasing peptide.";

RL Biochem. Biophys. Res. Commun. 178:526-530(1991).

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- SIMILARITY: BELONGS TO THE BOMBESIN/NEUROMEDIN B/RANATENSIN

CC FAMILY.

DR PIR: PQ0177; PQ0177.

DR InterPro; IPR000874; Bombesin.

DR Pfam; PF02044; Bombesin; 1.

DR PROSITE; PS00257; BOMBESIN; 1.

KW Bombesin family; Amidation.

FT MOD RES 10 10 AMIDATION.

SQ SEQUENCE 10 AA; 1094 MW; F81FBAE862CDC371 CRC64;

Query Match 33.3%; Score 2; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 2e+03; 0; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GH 2

DB 7 GH 8

RESULT 17

LSK2 LEUMA

ID LSK2 LEUMA STANDARD; PRT; 10 AA.

AC P09039;

DT 01-NOV-1988 (Rel. 09, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Leucosulfakinin-II (LSK-II).

OS Leucophaea maderae (Madeira cockroach), and

OS Periplaneta americana (American cockroach).

OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;

OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;

OC Blaberidae; Leucophaea.

OX NCBI_TaxID=6988, 6978;

RN [1]

RP SEQUENCE.

RX SPECIES=L.maderae;

RC MEDLINE=87048769; PubMed=3778455;

RA Nachman R.J., Holman G.M., Cook B.C., Haddor W.F., Ling N.;

RT "Leucosulfakinin-II, a blocked sulfated insect neuropeptide with

RT homology to cholecystokinin and gastrin.";

RL Biochem. Biophys. Res. Commun. 140:357-364(1986).

RN [2]

RP SEQUENCE.

RC SPECIES=P.americana; TISSUE=Corpora cardiaca;

RX MEDLINE=90137190; PubMed=2615921;

RA Veenstra J.A.;

RT "Isolation and structure of two gastrin/CCK-like neuropeptides from

RT the American cockroach homologous to the leucosulfakinins.";

RL Neuropeptides 14:145-149(1989).

CC -!- FUNCTION: CHANGE THE FREQUENCY AND AMPLITUDE OF CONTRACTIONS OF

CC THE COCKROACH HINGUT. STIMULATES MUSCLE CONTRACTION OF HINDGUT.

CC -!- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.

CC PIR: A26335; GMR02.

DR PIR: B60656; B60656.

DR InterPro; IPR001651; Gastrin.

DR PROSITE; PS00259; GASTRIN; 1.

KW Hormone; Amidation; Sulfation; Pyrrolidone carboxylic acid.

FT MOD RES 1 1 PYRROLIDONE CARBOXYLIC ACID.

FT MOD RES 5 5 SULFATION (IN L.MADERAE, BUT NOT IN

FT MOD RES 10 10 P.AMERICANA).

FT MOD RES 10 10 AMIDATION.

SQ SEQUENCE 10 AA; 1255 MW; 9B4F5391E86B5AAA CRC64;

Query Match

Best Local Similarity 100.0%; Pred. No. 2e+03; 0; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 1 GH 2
Db 6 GH 7

RESULT 18
THYL_PIG
ID THYL_PIG STANDARD; PRT; 3 AA.
AC P01151;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thyroliberin (Thyrotropin releasing hormone) (TRH) (Protirelin).
OS Sus acrofa (pig),
OS Ovis aries (Sheep),
OS Bombina orientalis (Oriental fire-bellied toad), and
OS Notophthalmus viridescens (Eastern newt) (Triturus viridescens).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823, 9940, 8346, 8316;
RN [1]
RN SEQUENCE.
RP SPECIES=Pig; TISSUE=Hypothalamus;
RC MEDLINE=70136150; PubMed=4984938;
RA Nair R.M.G., Barrett J.F., Bowers C.Y., Schally A.V.;
RT "Structure of porcine thyrotropin releasing hormone.";
RL Biochemistry 9:1103-1106(1970).
RN [2]
RN SYNTHESIS.
RP SPECIES=Pig;
RC MEDLINE=70039904; PubMed=4982117;
RA Boler J., Enzmann F., Folkers K., Bowers C.Y., Schally A.V.;
RT "The identity of chemical and hormonal properties of the thyrotropin
RT releasing hormone and pyroglutamyl-histidyl-proline amide.";
RL Biochem. Biophys. Res. Commun. 37:705-710(1969).
RN [3]
RN SEQUENCE.
RP SPECIES=Sheep; TISSUE=Hypothalamus;
RA Desiderio D.M. Jr., Burgess R., Dunn T.F., Vale W., Guillemin R.,
RA Ward D.N.;
RT "The elucidation of the primary structure of the hypothalamic thyroid
RT stimulating hormone releasing factor of ovine origin by means of mass
RT spectrometry.";
RL Org. Mass Spectrom. 5:221-228(1971).
RN [4]
RN SYNTHESIS.
RP SPECIES=Sheep;
RC MEDLINE=70163386; PubMed=4985794;
RA Burgess R., Dunn T.F., Desiderio D.M., Ward D.N., Vale W.,
RA Guillemin R.;
RT "Characterization of ovine hypothalamic hypophysiotropic
RT TSH-releasing factor.";
RL Nature 226:321-325(1970).
RN [5]
RN SEQUENCE.
RP SPECIES=B.orientalis; TISSUE=Skin;
RC MEDLINE=76138399; PubMed=815011;
RA Yasuhara T., Nakajima T.;
RT "Letter: Occurrence of Pyr-His-Pro-NH2 in the frog skin.";
RL Chem. Pharm. Bull. 23:3301-3303(1975).
RN [6]
RN SEQUENCE.
RP SPECIES=N.viridescens;
RC MEDLINE=75035605; PubMed=4214528;
RA Grimm-Joergensen Y., McKelvy J.P.;
RT "Biosynthesis of thyrotropin releasing factor by newt (Triturus
RT viridescens) brain in vitro. Isolation and characterization of
RT thyrotropin releasing factor.";
RL J. Neurochem. 23:471-478(1974).
CC -1- FUNCTION: TRH FUNCTIONS AS A REGULATOR OF THE BIOSYNTHESIS OF TSH
CC IN THE ANTERIOR PITUITARY GLAND AND AS A NEUROTRANSMITTER/
CC NEUROMODULATOR IN THE CENTRAL AND PERIPHERAL NERVOUS SYSTEMS.
DR PIR; A90919; RHDTDO.

DR PIR; A92971; A92971.
DR PIR; A93750; RBSHT.
KW Amidation; Pyrrolidone carboxylic acid.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 3 3 AMIDATION.
SQ SEQUENCE 3 AA; 380 MW; 7761F6B000000000 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.3e+05; Indels 0; Gaps 0;
Matches 1; Conservative 0; Mismatches 0;

Oy 2 H 2
Db 2 H 2

RESULT 19
ACHI_ACHFU
ID ACHI_ACHFU STANDARD; PRT; 4 AA.
AC P35904;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Achatina-I.
OS Achatina fulica (Giant African snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC Sigmurethra; Achatinoidea; Achatinidae; Achatina.
OX NCBI_TaxID=6530;
RN [1]
RN SEQUENCE, CHARACTERIZATION, AND SYNTHESIS.
RP STRAIN=Perussac; TISSUE=Ganglion;
RC MEDLINE=89273551; PubMed=2597281;
RA Kamatani Y., Minakata H., Kenny P.T.M., Iwashita T., Watanabe K.,
RA Funase K., Sun X.P., Yongsiri A., Kim K.H., Novales-Li P.,
RA Novales E.T., Kanapi C.G., Takeuchi H., Nomoto K.;
RT "Achatin-I, an endogenous neuroexcitatory tetrapeptide from Achatina
RT fulica ferussac containing a D-amino acid residue.";
RL Biochem. Biophys. Res. Commun. 160:1015-1020(1989).
RN [2]
RN CHARACTERIZATION.
RP STRAIN=Perussac; TISSUE=Heart atrium;
RC MEDLINE=91264856; PubMed=1675558;
RA Fujimoto K., Kubota I., Yasuda-Kamatani Y., Minakata H., Nomoto K.,
RA Yoshida M., Harada A., Muneoka Y., Kobayashi M.;
RT "Purification of achatin-I from the atria of the African giant snail,
RT Achatina fulica, and its possible function.";
RL Biochem. Biophys. Res. Commun. 177:847-853(1991).
RN [3]
RN X-RAY CRYSTALLOGRAPHY.
RP MEDLINE=93014529; PubMed=1399265;
RA Ishida T., In Y., Doi M., Inoue M., Yasuda-Kamatani Y., Minakata H.,
RA Iwashita T., Nomoto K.;
RT "Crystal structure and molecular conformation of achatin-I
RT (H-Gly-D-Phe-Ala-Asp-OH), an endogenous neuropeptide containing a
RT D-amino acid residue.";
RL Int. J. Pept. Protein Res. 39:258-264(1992).
CC -1- FUNCTION: NEUROEXCITATORY PEPTIDE; INCREASES THE IMPULSE FREQUENCY
CC AND PRODUCES A SPIKE BROADENING OF THE IDENTIFIED HEART EXCITATORY
CC NEURON (PON); ALSO ENHANCES THE AMPLITUDE AND FREQUENCY OF THE
CC HEART BEAT. HAS ALSO AN EFFECT ON SEVERAL OTHER MUSCLES.
DR PIR; A32480; A32480.
KW Hormone; D-amino acid.
FT MOD_RES 2 2
SQ SEQUENCE 4 AA; 408 MW; 6AADD9C810000000 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.3e+05; Indels 0; Gaps 0;
Matches 1; Conservative 0; Mismatches 0;

Oy 1 G 1
Db 1 G 1

```

```

RESULT 20
CPI OCTMI
ID OCPI OCTMI STANDARD; PRT; 4 AA.
AC P02731;
T 21-JUL-1986 (Rel. 01, Created)
T 21-JUL-1986 (Rel. 01, Last sequence update)
T 21-JUL-1986 (Rel. 01, Last annotation update)
T 21-JUL-1986 (Rel. 01, Last annotation update)
DE Eosinophilotoxic peptides.
DE Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE
MEDLINE=76078412; PubMed=1060093;
Gostzl E.J., Austen K.F.;
"Purification and synthesis of eosinophilotoxic tetrapeptides of
human lung tissue: identification as eosinophil chemotactic factor of
anaphylaxis.";
Proc. Natl. Acad. Sci. U.S.A. 72:4123-4127(1975).
C -1- MISCELLANEOUS: THESE PEPTIDES ARE RELEASED FROM MAST CELLS IN LUNG
(AND OTHER TISSUES) DURING HYPERSENSITIVITY REACTIONS
(ANAPHYLAXIS). THEIR ACTIVITIES, PREFERENTIALLY AFFECTING
EOSINOPHILS, INCLUDE CHEMOTAXIS, CHEMOTACTIC DEACTIVATION, RELEASE
OF ENZYMES, AND STIMULATION OF THE HEXOSE MONOPHOSPHATE SHUNT.
C GO: 00010105; P: anaphylaxis; IDA.
C GO: 0006935; P: chemotaxis; IDA.
C VARIANT 1 1 V -> A (IN OTHER PEPTIDE).
/FTID=VAR_005201.
SEQUENCE 4 AA; 390 MW; 6B05B862A0000000 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 G 1
b 2 G 2

RESULT 21
CPI OCTMI
ID OCPI OCTMI STANDARD; PRT; 4 AA.
C P58648;
T 28-FEB-2003 (Rel. 41, Created)
T 28-FEB-2003 (Rel. 41, Last sequence update)
T 28-FEB-2003 (Rel. 41, Last annotation update)
T 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cardioactive peptides Ocp-1/Ocp-2.
S Octopus minor (Octopus).
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;
OC Octopodiformes; Octopoda; Incirrata; Octopodidae; Octopus.
NCBI_TaxID=89766;
[1]
SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.
TISSUE=Brain;
MEDLINE=20336815; PubMed=10876044;
Iwakoshi E., Hisada M., Minakata H.;
"Cardioactive peptides isolated from the brain of a Japanese octopus,
Octopus minor.";
Peptides 21:623-630(2000).
C -1- FUNCTION: Cardioactive; has both positive chronotropic and
inotropic effects on the heart. Ocp-4 is a 1000 time less
active than Ocp-3.
C -1- SUBCELLULAR LOCATION: Secreted.
C -1- PTM: Ocp-4 has D-Ser instead of L-Ser.
C -1- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI.
KW Hormone; D-amino acid.
FT MOD_RES 2 2 D-SERINE (IN OCP-4).
SQ SEQUENCE 4 AA; 463 MW; 6AB365B810000000 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 G 1
b 2 G 2

RESULT 22
CPI OCTMI
ID OCPI OCTMI STANDARD; PRT; 4 AA.
AC P58649;
T 28-FEB-2003 (Rel. 41, Created)
T 28-FEB-2003 (Rel. 41, Last sequence update)
T 28-FEB-2003 (Rel. 41, Last annotation update)
T 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cardioactive peptides Ocp-3/Ocp-4.
S Octopus minor (Octopus).
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;
OC Octopodiformes; Octopoda; Incirrata; Octopodidae; Octopus.
NCBI_TaxID=89766;
[1]
SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.
TISSUE=Brain;
MEDLINE=20336815; PubMed=10876044;
Iwakoshi E., Hisada M., Minakata H.;
"Cardioactive peptides isolated from the brain of a Japanese octopus,
Octopus minor.";
Peptides 21:623-630(2000).
C -1- FUNCTION: Cardioactive; has both positive chronotropic and
inotropic effects on the heart. Ocp-4 is a 1000 time less
active than Ocp-3.
C -1- SUBCELLULAR LOCATION: Secreted.
C -1- PTM: Ocp-4 has D-Ser instead of L-Ser.
C -1- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI.
KW Hormone; D-amino acid.
FT MOD_RES 2 2 D-SERINE (IN OCP-4).
SQ SEQUENCE 4 AA; 463 MW; 6AB365B810000000 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 G 1
b 2 G 2

RESULT 23
CPI OCTMI
ID OCPI OCTMI STANDARD; PRT; 4 AA.
AC P36515;
T 01-JUN-1994 (Rel. 29, Created)
T 01-JUN-1994 (Rel. 29, Last sequence update)
T 01-JUN-1994 (Rel. 29, Last annotation update)
T 01-JUN-1994 (Rel. 29, Last annotation update)
DE Mitochondrial 60S ribosomal protein L1 (YmL1) (Fragment).
GN MRPL1.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
NCBI_TaxID=4932;
[1]
SEQUENCE.
MEDLINE=91285106; PubMed=2060626;
RX Grohmann L., Graack H.-R., Kruff V., Choli T., Goldschmidt-Reisin S.,
RA Kitakawa M.;
"Extended N-terminal sequencing of proteins of the large ribosomal
subunit from yeast mitochondria.";
FEBS Lett. 284:51-56(1991).
DR PIR; S17255; S17255.
DR SGD; L0002681; MRPL1.
KW Ribosomal protein; Mitochondrion.
FT NON_TER 4 4
SQ SEQUENCE 4 AA; 402 MW; 7771B2D5D000000000 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;

```

```

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
b 1 G 1

RESULT 22
CPI OCTMI
ID OCPI OCTMI STANDARD; PRT; 4 AA.
AC P58649;
T 28-FEB-2003 (Rel. 41, Created)
T 28-FEB-2003 (Rel. 41, Last sequence update)
T 28-FEB-2003 (Rel. 41, Last annotation update)
T 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cardioactive peptides Ocp-3/Ocp-4.
S Octopus minor (Octopus).
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;
OC Octopodiformes; Octopoda; Incirrata; Octopodidae; Octopus.
NCBI_TaxID=89766;
[1]
SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.
TISSUE=Brain;
MEDLINE=20336815; PubMed=10876044;
Iwakoshi E., Hisada M., Minakata H.;
"Cardioactive peptides isolated from the brain of a Japanese octopus,
Octopus minor.";
Peptides 21:623-630(2000).
C -1- FUNCTION: Cardioactive; has both positive chronotropic and
inotropic effects on the heart. Ocp-4 is a 1000 time less
active than Ocp-3.
C -1- SUBCELLULAR LOCATION: Secreted.
C -1- PTM: Ocp-4 has D-Ser instead of L-Ser.
C -1- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI.
KW Hormone; D-amino acid.
FT MOD_RES 2 2 D-SERINE (IN OCP-4).
SQ SEQUENCE 4 AA; 463 MW; 6AB365B810000000 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
b 1 G 1

RESULT 23
CPI OCTMI
ID OCPI OCTMI STANDARD; PRT; 4 AA.
AC P36515;
T 01-JUN-1994 (Rel. 29, Created)
T 01-JUN-1994 (Rel. 29, Last sequence update)
T 01-JUN-1994 (Rel. 29, Last annotation update)
T 01-JUN-1994 (Rel. 29, Last annotation update)
DE Mitochondrial 60S ribosomal protein L1 (YmL1) (Fragment).
GN MRPL1.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
NCBI_TaxID=4932;
[1]
SEQUENCE.
MEDLINE=91285106; PubMed=2060626;
RX Grohmann L., Graack H.-R., Kruff V., Choli T., Goldschmidt-Reisin S.,
RA Kitakawa M.;
"Extended N-terminal sequencing of proteins of the large ribosomal
subunit from yeast mitochondria.";
FEBS Lett. 284:51-56(1991).
DR PIR; S17255; S17255.
DR SGD; L0002681; MRPL1.
KW Ribosomal protein; Mitochondrion.
FT NON_TER 4 4
SQ SEQUENCE 4 AA; 402 MW; 7771B2D5D000000000 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
b 1 G 1

```

Query Match 16.7%; Score 1; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6
 |
 Db 1 S 1

RESULT 24
 AL14_CARMA STANDARD; PRT; 5 AA.
 ID E103 LITRU
 AC P82099;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Carcinus maenas 14.
 OS Carcinus maenas (Common shore crab) (Green crab).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
 OC Subrachyura; Portunoidae; Portunidae; Carcinus.
 OX NCBI_TaxID=6759;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
 RX MEDLINE=98121193; PubMed=9461295;
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
 RA Thorpe A.;
 RT "Isolation and identification of multiple neuropeptides of the
 RT allatostatin superfamily in the shore crab Carcinus maenas.";
 RL Eur. J. Biochem. 250:727-734(1997).
 CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 KW Neuropeptide; Amidation; Multigene family.
 FT MOD_RES 5 5
 FT SEQUENCE 5 AA; 586 MW; 672879D5AB300000 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6
 |
 Db 2 S 2

RESULT 25
 E103 LITRU STANDARD; PRT; 5 AA.
 ID E103 LITRU
 AC P82099;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Electrin 3.
 OS Litoria rubella (Desert tree frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
 OC Pelodyadinae; Litoria.
 OX NCBI_TaxID=104895;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Skin secretion;
 RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
 RT "Peptides from the skin glands of the Australian buzzing tree frog
 RT Litoria electrica. Comparison with the skin peptides from Litoria
 RT rubella.";
 RL Aust. J. Chem. 52:639-645(1999).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Skin.
 KW Amphibian defense peptide; Amidation.
 FT MOD_RES 5 5
 FT SEQUENCE 5 AA; 630 MW; 668761P2C9A00000 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 H 2
 |
 Db 3 H 3

RESULT 26
 E104 LITRU STANDARD; PRT; 5 AA.
 ID E104 LITRU
 AC P82100;
 DT 28-PEB-2003 (Rel. 41, Created)
 DT 28-PEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Electrin 4.
 OS Litoria rubella (Desert tree frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
 OC Pelodyadinae; Litoria.
 OX NCBI_TaxID=104895;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Skin secretion;
 RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
 RT "Peptides from the skin glands of the Australian buzzing tree frog
 RT Litoria electrica. Comparison with the skin peptides from Litoria
 RT rubella.";
 RL Aust. J. Chem. 52:639-645(1999).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Skin.
 KW Amphibian defense peptide; Amidation.
 FT MOD_RES 5 5
 FT SEQUENCE 5 AA; 616 MW; 61F2D1A059A00000 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 H 2
 |
 Db 5 H 5

RESULT 27
 PAP2_PARMA STANDARD; PRT; 5 AA.
 ID PAP2_PARMA
 AC P81864;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Pardaxin II (PXII) (Fragment).
 OS Pardachirus marmoratus (Red sea mores sole).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
 OC Soleiidae; Soleidae; Pardachirus.
 OX NCBI_TaxID=31087;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Skin secretion;
 RX MEDLINE=87057369; PubMed=3782138;
 RA Lazarovici P., Primor N., Loew L.M.;
 RT "Purification and pore-forming activity of two hydrophobic
 RT polypeptides from the secretion of the Red sea mores sole (Pardachirus
 RT marmoratus).";
 RL J. Biol. Chem. 261:16704-16713(1986).
 CC -!- FUNCTION: Exhibits unusual shark repellent and surfactant
 CC properties. Forms voltage-dependent, ion-permeable channels
 CC in membranes. At high concentration causes cell membrane lysis.
 CC -!- SUBUNIT: Monomer. In aqueous solution exists as a tetramer.
 CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- SIMILARITY: BELONGS TO THE PARDAXIN FAMILY.

KW Toxin.
FT NON_TER 5 5
SQ SEQUENCE 5 AA; 614 MW; 7769C9C9C8100000 CRC64;
Query Match 16.7%; Score 1; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 G 1
DB 1 G 1

RESULT 28

RE32_LITRU STANDARD; PRT; 5 AA.
ID RE32_LITRU
AC P82073;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Rubellidin 3.2.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Pelodyadinae; Litoria.
ON NCBI_TaxID=104895;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT "Peptides from the skin glands of the Australian buzzing tree frog
Litoria electrica. Comparison with the skin peptides from Litoria
rubella.";
RL Aust. J. Chem. 52:639-645(1999).
CC -!- FUNCTION: Shows neither neuropeptide activity nor antibiotic
activity.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.
KW Amphibian defense peptide.
SQ SEQUENCE 5 AA; 570 MW; 71A9C9C862A00000 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 G 1
DB 2 G 2

RESULT 29

TPIS CANFA STANDARD; PRT; 5 AA.
ID TPIS_CANFA
AC P54714;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Triosephosphate isomerase (EC 5.3.1.1) (TIM) (Fragment).
GN TP1.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
ON NCBI_TaxID=9615;
RN [1]
RP SEQUENCE
RC TISSUE=Heart;
RA MEDLINE=98163340; PubMed=9504812;
RX Dunn M.J., Corbett J.M., Wheeler C.H.;
RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
dog heart proteins.";
RL Electrophoresis 18:2795-2802(1997).
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate = glycerone

phosphate.
CC -!- PATHWAY: Plays an important role in several metabolic pathways.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SIMILARITY: BELONGS TO THE TRIOSEPHOSPHATE ISOMERASE FAMILY.
DR HSC-2DPAGE; P54714; DOG.
DR InterPro; IPR000652; Triophos_ismrse.
DR PROSITE; PS00171; TIM; PARTIAL.
KW Isomerase; Glycolysis; Gluconeogenesis; Fatty acid biosynthesis;
KW Pentose shunt.
FT NON_TER 1 1
FT NON_TER 5 5
SQ SEQUENCE 5 AA; 550 MW; 64444862C9A00000 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 G 1
DB 3 G 3

RESULT 30

UF01_MOUSE STANDARD; PRT; 5 AA.
ID UF01_MOUSE
AC P38639;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Unknown protein from 2D-page of fibroblasts (P19) (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RC TISSUE=Fibroblast;
RX MEDLINE=9500907; PubMed=7523108;
RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
RT "Separation and sequencing of familial and novel murine proteins
using preparative two-dimensional gel electrophoresis.";
RL Electrophoresis 15:735-745(1994).
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
PROTEIN IS: 6.6, ITS MW IS: 19 kDa.
FT NON_TER 5 5
SQ SEQUENCE 5 AA; 717 MW; 7364087043100000 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 G 1
DB 3 G 3

RESULT 31

UXA4_CHLTR STANDARD; PRT; 5 AA.
ID UXA4_CHLTR
AC P38005;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Unknown protein from 2D-page from elementary body (Fragment).
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
ON NCBI_TaxID=813;
RN [1]
RP SEQUENCE.
RC STRAIN=L2/434/Bu;
RA Bini L., Santucci A., Magi B., Marzocchi B., Sanchez-Campillo M.,
RA Comanducci M., Christianen G., Birkelund S., Vretou E., Ratti G.,
RA Pallini V.;

RL Submitted (SEP-1994) to the SWISS-PROT data bank.
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 4.5, ITS MW IS: 28 kDa.
 DR Siena-ZDPAGE; P38005; -.
 FT NON TER 5
 SQ SEQUENCE 5 AA; 474 MW; 75BAA865AA800000 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
 |
 Db 2 S 2

RESULT 32
 CIP1 MYTED STANDARD; PRT; 6 AA.
 ID CIP1 MYTED
 AC P13736;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Contraction-inhibiting peptide I (MIP I).
 OS Mytilus edulis (Blue mussel).
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloida;
 OC Mytiloidea; Mytilidae; Mytilus.
 OX NCBI_TaxID=6550;
 RN [1]
 RP SEQUENCE.
 RX TISSUE=Pedal ganglion;
 RC MEDLINE=89240357; PubMed=3377776;
 RA Hirata T., Kubota I., Iwasawa N., Takabatake I., Ikeda T., Munesaka Y.;
 RT "Structures and actions of Mytilus inhibitory peptides.";
 RL Biochem. Biophys. Res. Commun. 152:1376-1382(1988).
 CC -!- FUNCTION: INHIBITORY ACTION ON CONTRACTIONS IN SEVERAL MOLLUSCAN
 CC MUSCLES.
 CC -!- SIMILARITY: TO MIP II.
 DR PIR; A27696; A27696.
 KW Hormone; Amidation.
 FT MOD RES 6
 SQ SEQUENCE 6 AA; 637 MW; 72C9C68775B81000 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 |
 Db 1 G 1

RESULT 33
 CIP2 MYTED STANDARD; PRT; 6 AA.
 ID CIP2 MYTED
 AC P13737;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Contraction-inhibiting peptide II (MIP II).
 OS Mytilus edulis (Blue mussel).
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloida;
 OC Mytiloidea; Mytilidae; Mytilus.
 OX NCBI_TaxID=6550;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Pedal ganglion;
 RX MEDLINE=89240357; PubMed=3377776;
 RA Hirata T., Kubota I., Iwasawa N., Takabatake I., Ikeda T., Munesaka Y.;
 RT "Structures and actions of Mytilus inhibitory peptides.";
 RL Biochem. Biophys. Res. Commun. 152:1376-1382(1988).
 CC -!- FUNCTION: INHIBITORY ACTION ON CONTRACTIONS IN SEVERAL MOLLUSCAN
 CC MUSCLES.

CC -!- SIMILARITY: TO MIP I.
 DR PIR; B27696; B27696.
 KW Hormone; Amidation.
 FT MOD RES 6
 SQ SEQUENCE 6 AA; 621 MW; 72C9C6876DD81000 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 |
 Db 1 G 1

RESULT 34
 FARP MONEX STANDARD; PRT; 6 AA.
 ID FARP MONEX
 AC P41966;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE FMRamide-like neuropeptide GNFFRF-amide.
 OS Moniezia expansa (Sheep tapeworm).
 OC Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
 OC Cyclophyllidae; Anoplocephalidae; Moniezia.
 OX NCBI_TaxID=28841;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=93312289; PubMed=8323531;
 RA Maule A.G., Shaw C., Halton D.W., Thim L.;
 RT "GNFFRFamide: a novel FMRamide-immunoreactive peptide isolated from
 the sheep tapeworm, Moniezia expansa.";
 RL Biochem. Biophys. Res. Commun. 193:1054-1060(1993).
 CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
 CC FAMILY.
 DR PIR; A43129; A43129.
 KW Neuropeptide; Amidation.
 FT MOD RES 6
 SQ SEQUENCE 6 AA; 787 MW; 69D409C9C4481000 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 |
 Db 1 G 1

RESULT 35
 LOK1 LOCM1 STANDARD; PRT; 6 AA.
 ID LOK1 LOCM1
 AC P41491;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Locustakinin I.
 OS Locusta migratoria (Migratory locust).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;
 OC Acridoidea; Acrididae; Oedipodinae; Locusta.
 OX NCBI_TaxID=7004;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Corpora cardiaca;
 RX MEDLINE=92262851; PubMed=1585017;
 RA Schoofs L., Holman G.M., Proost P., van Damme J., Hayes T.K.,
 RA de Loof A.;
 RT "Locustakinin, a novel myotropic peptide from Locusta migratoria,
 RT isolation, primary structure and synthesis.";
 RL Regul. Pept. 37:49-57(1992).
 CC -!- FUNCTION: Myotropic peptide. May be important in the stimulation

CC of ion transport and inhibition of diuretic activity in Malpighian
tubules.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC PIR: A61068; A61068.

CC Neuropeptide; Amidation.

CC MOD_RES 6 6 AMIDATION.

CC SEQUENCE 6 AA; 654 MW; 686365A5B9CDB000 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 6;

Best Local Similarity 100.0%; Pred. No. 1.3e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 6 S 6

b 3 S 3

RESULT 36

TMOF_SARBU STANDARD; PRT; 6 AA.

C P41495;

T 01-NOV-1995 (Rel. 32, Created)

T 01-NOV-1995 (Rel. 32, Last sequence update)

T 01-FEB-1996 (Rel. 33, Last annotation update)

KE Trypsin-modulating oostatic factor (TMOF).

XC Sarcophaga bullata (Grey flesh fly) (Neobellieria bullata).

XC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

XC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroides;

XC Sarcophagidae; Sarcophaga.

NCBI_TaxID=7385;

[1]

P SEQUENCE, AND SYNTHESIS.

P TISSUE-Ovary;

MEDLIN=94211930; PubMed=8159807;

JA Bylens D., Borovsky D., Hunt D.F., Shabanowitz J., Grauwels L.,

de Loof A.;

T "Sequencing and characterization of trypsin modulating oostatic

factor (TMOF) from the ovaries of the grey fleshfly, Neobellieria

(Sarcophaga bullata)";

Regul. Pept. 50:61-72(1994).

-!- FUNCTION: HAS AN OOSTATIC ACTIVITY. INHIBITS TRYPSIN BIOSYNTHESIS

IN THE MIDGUT WHICH INDIRECTLY REDUCES THE VITELLOGENIN

CONCENTRATION IN THE HEMOLYPH RESULTING IN INHIBITION OF OOCYTE

DEVELOPMENT.

-!- DEVELOPMENTAL STAGE: SYNTHESIZED AND RELEASED FROM FOLLICULAR

CELLS.

Hormone.

SEQUENCE 6 AA; 695 MW; 61E72451B7642000 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 6;

Best Local Similarity 100.0%; Pred. No. 1.3e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 H 2

b 6 H 6

RESULT 37

RPI_PSEPU

D TRPI_PSEPU STANDARD; PRT; 6 AA.

C F36414;

T 01-JUN-1994 (Rel. 29, Created)

T 01-JUN-1994 (Rel. 29, Last sequence update)

T 01-FEB-1995 (Rel. 31, Last annotation update)

E TrpBA operon transcriptional activator (Fragment).

NCBI_TaxID=303;

[1]

P SEQUENCE FROM N.A.

S Pseudomonas putida.

C Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

C Pseudomonadaceae; Pseudomonas.

NCBI_TaxID=303;

[1]

P SEQUENCE FROM N.A.

RC STRAIN=PPG1 C15;
RX MEDLINE=89335826; PubMed=2503057;

RA Eberly L., Crawford I.P.;

RT "DNA sequence of the tryptophan synthase genes of Pseudomonas

putida";

RL Biochimie 71:521-531(1989).

CC -!- FUNCTION: ACTIVATES THE EXPRESSION OF THE TRPBA GENES ENCODING

THE TWO TRYPTOPHAN SYNTHASE SUBUNITS. IN THE ABSENCE OF THE

INDUCER (INDOLEGLYCEROL PHOSPHATE), TRPI BINDS UPSTREAM OF THE

TRPAB OPERON, OVERLAPPING ITS OWN PROMOTER REGION.

CC -!- SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL

REGULATORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

CC

DR EMBL; X13299; CAA31660.1; -.

DR InterPro; IPR000847; HTH_LYSR.

DR PROSITE; PS00044; HTH_LYSR_FAMILY; PARTIAL.

KW Tryptophan biosynthesis; Transcription regulation; Activator;

KW DNA-binding.

FT NON_TER 6 6

SQ SEQUENCE 6 AA; 693 MW; 77672AA1EDD6F000 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 6;

Best Local Similarity 100.0%; Pred. No. 1.3e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 H 2

Db 3 H 3

RESULT 38

ALL2_CARMA

ID _ALL2_CARMA STANDARD; PRT; 7 AA.

AC F81805;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Carcinostat 2.

OS Carcinus maenas (Common shore crab) (Green crab).

OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;

OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;

OC Eubrachyura; Portunoidae; Portunidae; Carcinus.

NCBI_TaxID=6759;

[1]

RN

SEQUENCE.

RC TISSUE=Cerebral ganglion, and Thoracic ganglion;

RX MEDLINE=98121193; PubMed=9461295;

RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,

Thorpe A.;

RT "Isolation and identification of multiple neuropeptides of the

allatostatin superfamily in the shore crab Carcinus maenas.";

Eur. J. Biochem. 250:727-734(1997).

CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.

CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.

KW Neuropeptide; Amidation; Multigene family.

FT MOD_RES 7 7

SQ SEQUENCE 7 AA; 770 MW; 672879DCB5DB70 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 7;

Best Local Similarity 100.0%; Pred. No. 1.3e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1

Db 6 G 6

RESULT 39
ALL3 CARMA
ID ALL3 CARMA STANDARD; PRT; 7 AA.
AC P81806;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 3.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Multigene family.
FT MOD RES 7
SQ SEQUENCE 7 AA; 796 MW; 672879CDBC476B70 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
DB 6 G 6

RESULT 40
ALL4 CARMA
ID ALL4 CARMA STANDARD; PRT; 7 AA.
AC P81807;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 4.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Multigene family.
FT MOD RES 7
SQ SEQUENCE 7 AA; 782 MW; 672879CDBC476AC0 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
DB 6 G 6

RESULT 41
ALL5 CARMA
ID ALL5 CARMA STANDARD; PRT; 7 AA.
AC P81808;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 5.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Multigene family.
FT MOD RES 7
SQ SEQUENCE 7 AA; 781 MW; 672879CDBC476420 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
DB 6 G 6

RESULT 42
ALL7 CYDPO
ID ALL7 CYDPO STANDARD; PRT; 7 AA.
AC P82158;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cydiastatin 7.
OS Cydia pomonella (Codling moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Tortricoidae; Tortricidae; Olethreutinae; Cydia.
OX NCBI_TaxID=82606;
RN [1]
RP SEQUENCE.
RC TISSUE=Larva;
RX MEDLINE=98054539; PubMed=9192829;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D.,
RA Davey M., East P.D., Thorpe A.;
RT "Lepidopteran peptides of the allatostatin superfamily.";
RL Peptides 18:1301-1309(1997).
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation.
FT MOD RES 7
SQ SEQUENCE 7 AA; 873 MW; 672879CABB569350 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
DB 6 G 6


```

RESULT 43
BRHP CONTM
ID BRHP CONTM STANDARD; PRT; 7 AA.
AC P58603;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Bromheptapeptide Im.
OS Conus imperialis (Imperial cone).
OS Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=35631;
[1]
RN SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.
RP TISSUE=Venom;
RC MEDLINE=97184108; PubMed=9030520;
RA Craig A.G., Jimenez E.C., Dykert J., Nielsen D.B., Gulyas J.,
RA Abogadie F.C., Porter J., Rivier J.E., Cruz L.J., Olivera B.M.,
RA McIntosh J.M.;
RT "A novel post-translational modification involving bromination of
RT tryptophan. Identification of the residue, L-6-bromotryptophan, in
RT peptides from Conus imperialis and Conus radiatus venom.";
RL J. Biol. Chem. 272:4689-4698(1997).
CC -!- FUNCTION: Does not elicit gross behavioral symptoms when injected
CC centrally or peripherally in mice.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -!- MASS SPECTROMETRY: MW=853.19; METHOD=LSIMS.
CC PIR; A58512;
DR Bromination; Amidation; Pyrrolidone carboxylic acid.
KW DISULFID 2 7
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 6 6 BROMINATION.
FT MOD_RES 7 7 AMIDATION.
SQ SEQUENCE 7 AA; 795 MW; 6EA37DC6D87EA6B0 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
DB 3 G 3

RESULT 44
CIA_ENTFA
ID CIA_ENTFA STANDARD; PRT; 7 AA.
AC P11932;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Sex pheromone CAM373 (Clumping-inducing agent) (CIA).
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;
[1]
RN SEQUENCE.
RP MEDLINE=87005252; PubMed=3093276;
RA Mori M., Tanaka H., Sakagami Y., Isogai A., Fujino M., Kitada C.,
RA White B.A., An F.Y., Clewell D.B., Suzuki A.;
RT "Isolation and structure of the streptococcus faecalis sex pheromone,
RT CAM373."
RT FEBS Lett. 206:69-72(1986).
CC -!- FUNCTION: CAM373 INDUCES MATING RESPONSE OF DONOR CELLS
CC HARBORING PAM373.
CC -!- MISCELLANEOUS: THE N-TERMINUS IS POSSIBLY RESPONSIBLE FOR
CC SPECIFICITY OF PHEROMONES TO PLASMIDS.
CC -!- SIMILARITY: C-TERMINAL TO THE PHEROMONES CPD1 AND CAD1.
DR PIR; A25269; A25269.
KW Pheromone.
SQ SEQUENCE 7 AA; 734 MW; 75BDD72059C05DB0 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
DB 3 G 3

RESULT 45
FAR1 HELTI
ID FAR1 HELTI STANDARD; PRT; 7 AA.
AC P41871;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE FMRFamide-like neuropeptide GDPFLRF-amide.
OS Helisoma trivolvis (Snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Bascommatophora;
OC Lymnaeidae; Planorbidae; Helisoma.
OX NCBI_TaxID=27815;
[1]
RN SEQUENCE.
RP TISSUE=Kidney;
RC MEDLINE=94286417; PubMed=7912428;
RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;
RT "FMRFamide-related peptides from the kidney of the snail, Helisoma
RT trivolvis.";
RL Peptides 15:31-36(1994).
CC -!- FUNCTION: APPEARS TO BE INVOLVED IN OSMOREGULATION BY AFFECTING
CC THE KIDNEY, MANTLE AND SKIN.
CC -!- TISSUE SPECIFICITY: KIDNEY, SKIN, MANTLE AND THE HEMOLYMPH.
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 7 7
SQ SEQUENCE 7 AA; 851 MW; 69D40729D76AAB10 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
DB 1 G 1

RESULT 46
FAR2 ASCSU
ID FAR2 ASCSU STANDARD; PRT; 7 AA.
AC P31890;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE FMRFamide-like neuropeptide Af2.
OS Ascaris suum (Pig roundworm) (Ascaris lumbricoides), and
OS Panagrellus redivivus.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC Ascarididae; Ascaris.
OX NCBI_TaxID=6253, 6233;
[1]
RN SEQUENCE.
RP SPECIES=A.suum;
RC MEDLINE=9332431; PubMed=8332542;
RA Cowden C., Stretton A.O.W.;
RT "Af2, an Ascaris neuropeptide: isolation, sequence, and bioactivity.";
RL Peptides 14:423-430(1993).
[2]
RN SEQUENCE.
RP SPECIES=P.redivivus;
RC MEDLINE=95060998; PubMed=7970891;
RA Maule A.G., Shaw C., Bowman J.W.;

```

```

RT "The FMRamide-like neuropeptide AF2 (Ascaris suum) is present in the
RT free-living nematode, Panagrellus redivivus (Nematoda, Rhabditida).";
RL Parasitology 109:351-356(1994).
CC -!- FUNCTION: HAS EFFECTS ON MUSCLE TENSION.
CC -!- TISSUE SPECIFICITY: FOUND IN THE NERVE CORDS AND A VARIETY OF
CC -!- GANGLIA PARTICULARLY IN THE ANTERIOR REGIONS.
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD RES 7 7 AMIDATION.
SQ SEQUENCE 7 AA; 992 MW; 69D4073B5B11E350 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 H 2
Db 2 H 2

RESULT 47
FAR3 HAECO STANDARD; PRT; 7 AA.
AC P81258;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE FMRamide-like neuropeptide PF3 (KSAYMRF-amide).
OS Haemonchus contortus (Barber pole worm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Trichostrongyloidea; Haemonchidae; Haemonchinae; Haemonchus.
OX NCBI_TaxID=6289;
RN [1]
RP TISSUE=Neuron;
RC MEDLINE=99318264; PubMed=10391380;
RA Marks N.J., Sangster N.C., Maule A.G., Halton D.W., Thompson D.P.,
RA Geary T.G., Shaw C.;
RT "Structural characterization and pharmacology of KHEVLRamide (AF2)
RT and KSAYMRFamide (PF3/AF8) from Haemonchus contortus.";
RL Mol. Biochem. Parasitol. 100:185-194(1999).
CC -!- FUNCTION: ACTIVE ON NEUROMUSCULATURE.
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD RES 7 7 AMIDATION.
SQ SEQUENCE 7 AA; 902 MW; 69D4068B5DC5B350 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6
Db 2 S 2

RESULT 48
FAR3 PANRE STANDARD; PRT; 7 AA.
AC P41874;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE FMRamide-like neuropeptide PF3 (KSAYMRF-amide).
OS Panagrellus redivivus.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
OC Panagrolaimoidea; Panagrolaimidae; Panagrellus.
OX NCBI_TaxID=6233;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC MEDLINE=94235053; PubMed=8179635;

```

```

RA Maule A.G., Shaw C., Bowman J.W., Halton D.W., Thompson D.P.,
RA Geary T.G., Thim L.;
RT "KSAYMRFamide: a novel FMRamide-related heptapeptide from the free-
RT living nematode, Panagrellus redivivus, which is myoactive in the
RT parasitic nematode, Ascaris suum.";
RL Biochem. Biophys. Res. Commun. 200:973-980(1994).
CC -!- FUNCTION: MYOACTIVE; INDUCES A RAPID CONCENTRATION-DEPENDENT
CC MUSCLE TENSION INCREASE.
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
DR PIR; PC2132; PC2132.
KW Neuropeptide; Amidation.
FT MOD RES 7 7 AMIDATION.
SQ SEQUENCE 7 AA; 902 MW; 69D4068B5DC5B350 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6
Db 2 S 2

RESULT 49
FAR5 HIRME STANDARD; PRT; 7 AA.
AC P42564;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE FMRamide-like neuropeptide GKYMRF-amide.
OS Hirudo medicinalis (Medicinal leech).
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC Arynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
OX NCBI_TaxID=6421;
RN [1]
RP SEQUENCE.
RX MEDLINE=92195954; PubMed=1686933;
RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT "Identification of RFamide neuropeptides in the medicinal leech.";
RL Peptides 12:897-908(1991).
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD RES 7 7 AMIDATION.
SQ SEQUENCE 7 AA; 858 MW; 69D4068B53387810 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 G 1
Db 1 G 1

RESULT 50
GFRP MOUSE STANDARD; PRT; 7 AA.
ID GFRP MOUSE STANDARD; PRT; 7 AA.
AC P99025;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE GTP cyclohydrolase I feedback regulatory protein (P35) (Fragment).
GN GCHFR OR GFRP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RC TISSUE=Liver;

```

RA Sanchez J.-C., Rouge V., Frutiger S., Hughes G., Yan J.X.,
 RA Hoogland C., Appel R.D., Binz P.-A., Hochstrasser D.F.,
 RA Cowthorne M.,
 RL Submitted (AUG-1998) to the SWISS-PROT data bank.
 CC -!- FUNCTION: MEDIATES TETRAHYDROBIOTERIN INHIBITION OF GTP
 CC CYCLOHYDROLASE 1. THIS INHIBITION IS REVERSED BY L-PHENYLALANINE
 CC (BY SIMILARITY).
 CC -!- SUBUNIT: Homodimer (By similarity).
 CC SWISS-2DPAGE: P99025; MOUSE.
 DR INIT MET 0 0
 FT NON TER 7 7
 FT SEQUENCE 7 AA; 806 MW; 71BSB057273B4700 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
 |
 Db 6 S 6

RESULT 51
 IGAO DACDE STANDARD; PRT; 7 AA.
 ID TGAO DACDE
 AC P06294;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)
 DE Galactose oxidase inhibitor.
 OS Dactylium dendroides (Cladobotryum dendroides).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypomyces.
 OX NCBI_TaxID=5132;
 RN [1]
 RP SEQUENCE.
 RA Avidig G., Markus Z.;
 RT "Identification of a peptide inhibitor of galactose oxidase from
 RT Dactylium dendroides".
 RL Fed. Proc. 31:447-447(1972).
 CC -!- FUNCTION: BINDS ONE COPPER ION PER MOLECULE BUT DOES NOT BIND THE
 CC GALACTOSE OXIDASE APOENZYME. IT MAY INACTIVATE THE ENZYME BY
 CC BINDING TO ITS PROSTHETIC COPPER GROUP.
 DR PIR; A01341; XEYDGP.
 KW Copper; Metalloenzyme inhibitor.
 SQ SEQUENCE 7 AA; 706 MW; 75BB01A456D87D80 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 |
 Db 2 G 2

RESULT 52
 LANC CARUI STANDARD; PRT; 7 AA.
 ID LANC CARUI
 AC P36960;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lantibiotic carnocin U149 (Fragment).
 OS Carnobacterium sp. (strain U149).
 CC Bacteria; Firmicutes; Lactobacillales; Carnobacteriaceae;
 CC Carnobacterium.
 OX NCBI_TaxID=35782;
 RN [1]
 RP SEQUENCE.
 RA Stoffels G., Nissen-Meyer J., Gudmundsdottir A., Sletten K., Holo H.,
 RA Nes I.F.;

RT "Purification and characterization of a new bacteriocin isolated from
 RT a Carnobacterium sp.";
 RL Appl. Environ. Microbiol. 58:1417-1422(1992).
 CC -!- FUNCTION: LANTHIONINE-CONTAINING PEPTIDE ANTIBIOTIC (LANTIBIOTIC).
 CC ACTIVE ON GRAM-POSITIVE BACTERIA.
 KW Antibiotic; Bacteriocin; Lantibiotic.
 FT NON TER 7 7
 FT SEQUENCE 7 AA; 786 MW; 741776D05805B810 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 |
 Db 1 G 1

RESULT 53
 MNPI LEPDE STANDARD; PRT; 7 AA.
 ID MNPI LEPDE
 AC P42984;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Myotropic neuropeptide 1 (Led-MNP-I).
 OS Leptinotarsa decemlineata (Colorado potato beetle).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
 OC Phytophaga; Chrysomeloidea; Chrysomelidae; Chrysomelinae;
 OC Chrysomelini; Leptinotarsa.
 OX NCBI_TaxID=7539;
 RN [1]
 RP SEQUENCE, AND SYNTHESIS.
 RC TISSUE=Head;
 RX MEDLINE=95380343; PubMed=7651886;
 RA Spittaels L., Vankeerberghen A., Schoofs L., Torrekens S.,
 RA Grauwels L., van Leuven F., de Loof A.;
 RT "Identification, characterization, and immunological localization of
 RT a novel myotropic neuropeptide in the Colorado potato beetle.
 RT Leptinotarsa decemlineata".
 RL Peptides 16:365-374(1995).
 CC -!- FUNCTION: MYOTROPIC PEPTIDE. STIMULATES THE CONTRACTIONS OF THE
 CC OVIDUCT.
 KW Neuropeptide; Amidation.
 FT MOD RES 7 7
 FT SEQUENCE 7 AA; 705 MW; 6DD73768745B5DB0 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 |
 Db 4 G 4

RESULT 54
 PPH2 LYCES STANDARD; PRT; 7 AA.
 ID PPH2 LYCES
 AC P83379;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Purple acid phosphatase isozyme LesAP2 (EC 3.1.3.2) (fragment).
 OS Lycopersicon esculentum (tomato).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Asteridae; lamids; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4081;
 RN [1]
 RP SEQUENCE, CATALYTIC ACTIVITY, SUBUNIT, SUBCELLULAR LOCATION, AND
 RP GLYCOSYLATION.

RC STRAIN=cv, MoneyMaker; TISSUE=Seed;
 RX MEDLINE=22361242; PubMed=12473124;
 RA Bozzo G.G.; Raghothama K.G.; Plaxton W.C.;
 RT "Purification and characterization of two secreted purple acid
 phosphatase isozymes from phosphate-starved tomato (*Lycopersicon
 esculentum*) cell cultures.";
 RL Eur. J. Biochem. 269:6278-6286 (2002).
 CC -!- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an
 CC alcohol + phosphate.
 CC -!- SUBUNIT: Monomer.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- PTM: Glycosylated.
 CC -!- MISCELLANEOUS: In *L. esculentum* there are at least two isozymes of
 CC purple acid phosphatase.
 CC Hydrolase; Glycoprotein.
 FT NON_TER 1 7
 FT NON_TER 1 7
 SQ SEQUENCE 7 AA; 810 MW; 672AA962C9C729A0 CRC64;
 Query Match 16.7%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 G 1
 Db 5 G 5
 RESULT 55
 TYS1_LITRU STANDARD; PRT; 7 AA.
 AC P82065;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Tryptophyllin 5.1.
 OS *Litoria rubella* (Desert tree frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoides; Hylidae;
 OC Pelodyadinae; Litoria.
 OX NCBI_TaxID=104895;
 RN [1]
 RP SEQUENCE, AND MASS SPECTROMETRY.
 RC TISSUE=Skin secretion;
 RA Steinborner S.T.; Wabnitz P.A.; Waugh R.J.; Bowie J.H.; Gao C.;
 RA Tyler M.J.; Wallace J.C.;
 RT "The structure of new peptides from the Australian red tree frog
 RT '*Litoria rubella*'. The skin peptide profile as a probe for the study
 RT of evolutionary trends of amphibians.";
 RL Aust. J. Chem. 49:955-963 (1996).
 CC -!- FUNCTION: May act as a neuromodulator or neurotransmitter.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.
 CC -!- MASS SPECTROMETRY: MW=965; METHOD=FAB.
 KW Amphibian defense peptide; Amidation; Neuropeptide;
 KW Pyrrolidone carboxylic acid. PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 1 1
 FT MOD_RES 7 7
 FT MOD_RES 7 7
 SQ SEQUENCE 7 AA; 983 MW; 7401E9D3676046B0 CRC64;
 Query Match 16.7%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 H 2
 Db 6 H 6
 RESULT 56
 UC24_MAIZE STANDARD; PRT; 7 AA.
 ID UC24_MAIZE
 AC P80630;

DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Unknown protein from 2D-page of etiolated coleoptile (Spot 447)
 DE (Fragment).
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACAD clade; Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Coleoptile;
 RA Touzet P.; Riccardi F.; Morin C.; Damerval C.; Huet J.-C.;
 RA Pernollet J.-C.; Zivy M.; de Vienne D.;
 RT "The maize two dimensional gel protein database: towards an integrated
 RT genome analysis program.";
 RL Theor. Appl. Genet. 93:997-1005 (1996).
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED P1 OF THIS UNKNOWN
 CC PROTEIN IS: 6.0. ITS MW IS: 30.0 kDa.
 CC Maize-2DPAGE; P80630; COLSOPHILS.
 DR MaizeDB; 123956; --
 FT NON_TER 1 1
 FT NON_TER 7 7
 SQ SEQUENCE 7 AA; 665 MW; 6DC1B5B33DC1B5D0 CRC64;
 Query Match 16.7%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 6 S 6
 Db 1 S 1
 RESULT 57
 UF03_MOUSE STANDARD; PRT; 7 AA.
 ID UF03_MOUSE
 AC P38641;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE Unknown protein from 2D-page of fibroblasts (P36) (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Fibroblast;
 RX MEDLINE=95009907; PubMed=7523108;
 RA Merrick B.A.; Patterson R.M.; Wichter L.L.; He C.; Selkirk J.K.;
 RT "Separation and sequencing of familial and novel murine proteins
 RT using preparative two-dimensional gel electrophoresis.";
 RL Electrophoresis 15:735-745 (1994).
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED P1 OF THIS UNKNOWN
 CC PROTEIN IS: 5.1. ITS MW IS: 36 kDa.
 FT NON_TER 1 1
 FT NON_TER 7 7
 SQ SEQUENCE 7 AA; 842 MW; 6AA72B1DDB1B1180 CRC64;
 Query Match 16.7%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 H 2
 Db 1 H 1
 RESULT 58
 UF04_MOUSE STANDARD; PRT; 7 AA.
 ID UF04_MOUSE
 AC P38642;

DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE Unknown protein from 2D-page of fibroblasts (P46) (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 [1]
 RP SEQUENCE.
 RC TISSUE=Fibroblast;
 RX MEDLINE=95009307; PubMed=7523108;
 RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
 RA "Separation and sequencing of familial and novel murine proteins
 RT using preparative two-dimensional gel electrophoresis.";
 RL Electrophoresis 15:735-745(1994).
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 5.0, ITS MW IS: 46 kDa.
 CC NON TER 7
 CC SEQUENCE 7 AA; 766 MW; 68640AB777632700 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y 1 G 1

Db 7 G 7

RESULT 59

JH11 RAT

ID UH11 RAT

AC P56576;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Unknown protein from 2D-page of heart tissue (Spot P11) (Fragment).

DS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

NCBI_TaxID=10116;

[1]
 RP SEQUENCE.
 RC STRAIN=Wistar; TISSUE=Heart;
 RA Li X.-P., Pleissner K.-P., Scheler C., Regitz-Zagrosek V., Salikov J.,
 RA Submitted (SEP-1998) to the SWISS-PROT data bank.
 LA Jungblut P.R.;
 L -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 C PROTEIN IS: 8.5, ITS MW IS: 42 kDa.
 C UNSURE 2 2
 C OR A.
 C NON TER 7 7
 C SEQUENCE 7 AA; 775 MW; 6866DB040DC5A6B0 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y 6 S 6

Db 2 S 2

RESULT 60

IN06 PINPS

D UN06 PINPS

AC P81675;

DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Unknown protein from 2D-page of needles (N141) (Fragment).

OS Pinus pinaster (Maritime pine).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
 NCBI_TaxID=71647;
 [1]
 RP SEQUENCE.
 RC TISSUE=Needle;
 RX MEDLINE=99274088; PubMed=10344291;
 RA Costa P., Pionneau C., Bauw G., Dubos C., Bahrman N., Kremer A.,
 RA Frigerio J.-M., Plomion C.;
 RA "Separation and characterization of needle and xylem maritime pine
 RT proteins.";
 RL Electrophoresis 20:1098-1108(1999).
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 6.6, ITS MW IS: 25 kDa.
 CC NON TER 1 1
 CC NON TER 7 7
 CC SEQUENCE 7 AA; 823 MW; 69D76724486B5740 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1

Db 3 G 3

RESULT 61

WWA1 ACHFU

ID WWA1 ACHFU

AC P35919;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 01-OCT-1994 (Rel. 30, Last annotation update)

DE Wamide-1.

OS Achatina fulica (Giant African snail).

OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;

OC Sigmurethra; Achatinoidea; Achatinidae; Achatina.

NCBI_TaxID=6530;

[1]
 RP SEQUENCE.
 RC TISSUE=Ganglion;
 RA MEDLINE=93265912; PubMed=8495720;
 RA Minakata H., Ikeda T., Muneoka Y., Kobayashi M., Nomoto K.;
 RA "Wamide-1, -2 and -3: novel neuromodulatory peptides isolated from
 RT ganglia of the African giant snail, Achatina fulica.";
 RL FEBS Lett. 323:104-108(1993).
 CC -1- FUNCTION: EXHIBITS MODULATORY EFFECTS ON THE PERIPHERAL NERVOUS
 CC SYSTEM. INHIBITS ACTIVITY ON A CENTRAL NEURON.
 CC PIR; S33245; S33245.
 CC Neuropeptide; Amidation.
 CC MOD RES 7 7
 CC SEQUENCE 7 AA; 993 MW; 7362D5B6B041310 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6

Db 5 S 5

RESULT 62

WWA2 ACHFU

ID WWA2 ACHFU

AC P35920;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 01-OCT-1994 (Rel. 30, Last annotation update)

DE Wamide-2.

OS Achatina fulica (Giant African snail).

OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;

OC Sigmurethra; Achatinoidea; Achatinidae; Achatina.

```

OX NCBI_TaxID=6530;
RN [1]
RP SEQUENCE.
RC TISSUE=Ganglion;
RX MEDLINE=933265912; PubMed=8495720;
RA Minakata H., Ikeda T., Muneoka Y., Kobayashi M., Nomoto K.;
RT "Wamide-1, -2 and -3: novel neuromodulatory peptides isolated from
RL ganglia of the African giant snail, Achatina fulica.";
DR FEBS Lett. 323:104-108(1993).
DR PIR: S33246; S33246.
KW Neuropeptide; Amidation.
FT MOD_RES 7 7
SQ SEQUENCE 7 AA; 964 MW; 7362D5B686D32310 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6
Db 5 S 5

RESULT 63
WMA3_ACHFU STANDARD; PRT; 7 AA.
AC P35921;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE "Wamide-3.
OS Achatina fulica (Giant African snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC Sigmurethra; Achatinoidea; Achatinidae; Achatina.
OX NCBI_TaxID=6530;
RN [1]
RP SEQUENCE.
RC TISSUE=Ganglion;
RX MEDLINE=933265912; PubMed=8495720;
RA Minakata H., Ikeda T., Muneoka Y., Kobayashi M., Nomoto K.;
RT "Wamide-1, -2 and -3: novel neuromodulatory peptides isolated from
RL ganglia of the African giant snail, Achatina fulica.";
DR FEBS Lett. 323:104-108(1993).
DR PIR: S33244; S33244.
KW Neuropeptide; Amidation.
FT MOD_RES 7 7
SQ SEQUENCE 7 AA; 965 MW; 7362D5B69B132310 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6
Db 5 S 5

RESULT 64
ACI_THUAL STANDARD; PRT; 8 AA.
AC P18591;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-NOV-1990 (Rel. 16, Last annotation update)
DE Angiotensin-converting enzyme inhibitor.
OS Thunus albacares (Yellowfin tuna) (Neothunnus macropterus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
OC Scombridae; Thunnus.
OX NCBI_TaxID=8236;
RN [1]
RP SEQUENCE.

```

```

RC TISSUE=Muscle;
RX MEDLINE=88326322; PubMed=3415688;
RA Kohama Y., Matsumoto S., Oka H., Teramoto T., Okabe M., Minura T.;
RT "Isolation of angiotensin-converting enzyme inhibitor from tuna
RL muscle.";
DR PIR: A31570; A31570.
SQ SEQUENCE 8 AA; 953 MW; 6AA863733051F1B7 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 H 2
Db 3 H 3

RESULT 65
AKHG_GRYBI STANDARD; PRT; 8 AA.
AC P14086;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Adipokinetic hormone G (AKH-G) (RO II).
OS Gryllus bimaculatus (Two-spotted cricket), and
OS Romalea microptera (lubber grasshopper).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Orthoptera; Ensifera; Gryllidae; Gryllinae;
OC Gryllus.
OX NCBI_TaxID=6999, 7007;
RN [1]
RP SEQUENCE.
RC SPECIES=G.bimaculatus; TISSUE=Corpora cardiaca;
RX MEDLINE=88106553; PubMed=3426616;
RA Gaede G., Rinehart K.L. Jr.;
RT "Primary sequence analysis by fast atom bombardment mass spectrometry
of a peptide with adipokinetic activity from the corpora cardiaca of
the cricket Gryllus bimaculatus.";
RL Biochem. Biophys. Res. Commun. 149:908-914(1987).
RN [2]
RP SEQUENCE.
RC SPECIES=R.microptera; TISSUE=Corpora cardiaca;
RX MEDLINE=89145002; PubMed=3226948;
RA Gaede G., Hilbich C., Beyreuther K., Rinehart K.L. Jr.;
RT "Sequence analyses of two neuropeptides of the AKH/RPCH-family from
the lubber grasshopper, Romalea microptera.";
RL Peptides 9:681-688(1988).
CC -!- FUNCTION: THIS HORMONE, RELEASED FROM CELLS IN THE CORPORA
CARDIACA AFTER THE BEGINNING OF FLIGHT, CAUSES RELEASE OF
DIGLYCERIDES FROM THE FAT BODY AND THEN STIMULATES THE FLIGHT
MUSCLES TO USE THESE DIGLYCERIDES AS AN ENERGY SOURCE.
CC -!- SIMILARITY: BELONGS TO THE AKH / RHTH / RPCH FAMILY.
DR PIR: A28004; A28004.
DR InterPro: IPR02047; AKH.
RC PROSITE: PS00256; AKH; 1.
KW Neuropeptide; Amidation; Flight; Pyrrolidone carboxylic acid.
FT MOD_RES 1 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 8 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 938 MW; 867861B5B9C452D6 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6
Db 5 S 5

RESULT 66
AKH_L1BAU

```

ID AKH_LIBAU STANDARD; PRT; 8 AA.
AC P25418;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Adipokinetic hormone (AKH)
OS Libellula auripennis (Skimmer dragonfly).
OS Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Palaeoptera; Odonata; Anisoptera; Libellulidae; Libellula.
OX NCBI_TaxID=6966;
RN [1]
RP SEQUENCE AND SYNTHESIS.
RC TISSUE=Corpora cardiaca;
RX MEDLINE=90359055; PubMed=2390213;
RA Gaede G.;
RT "The putative ancestral peptide of the adipokinetic/red-pigment-
RT concentrating hormone family isolated and sequenced from a
RT dragonfly.";
RL Biol. Chem. Hoppe-Seyler 371:475-483(1990).
CC -!- FUNCTION: THIS HORMONE, RELEASED FROM CELLS IN THE CORPORA
CC CARDIACA AFTER THE BEGINNING OF FLIGHT, CAUSES RELEASE OF
CC DIGLYCERIDES FROM THE FAT BODY AND THEN STIMULATES THE FLIGHT
CC MUSCLES TO USE THESE DIGLYCERIDES AS AN ENERGY SOURCE.
CC -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR PIR; S10596; S10596.
DR InterPro; IPR002047; AKH.
DR PROSITE; PS00256; AKH; 1.
KW Neuropeptide; Amidation; Flight; Pyrrolidone carboxylic acid.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 978 MW; 8665A771A9C452D6 CRC64;
Query Match 16.7%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 6 S 6
Db 7 S 7

RESULT 67
AKH_MELML
ID AKH_MELML STANDARD; PRT; 8 AA.
AC P25423;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Adipokinetic hormone (AKH)
OS Melolontha melolontha (Cockchafer),
OS Geotrupes stercorosus (Dor beetle), and
OS Pachnoda marginata (Flower beetle).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Scarabaeiformia;
OC Scarabaeidae; Melolonthinae; Melolontha.
OX NCBI_TaxID=7061, 7087, 7058;
RN [1]
RP SEQUENCE.
RC SPECIES=M. melolontha, and G. stercorosus; TISSUE=Corpora cardiaca;
RX MEDLINE=91248100; PubMed=2039445;
RA Gaede G.;
RT "A unique charged tyrosine-containing member of the adipokinetic
RT hormone/red-pigment-concentrating hormone peptide family isolated and
RT sequenced from two beetle species.";
RL Biochem. J. 275:671-677(1991).
RN [2]
RP SEQUENCE.
RC SPECIES=P. marginata; TISSUE=Corpora cardiaca;
RX MEDLINE=92265187; PubMed=1586453;
RA Gaede G., Lopata A., Kellner R., Rinehart K.L. Jr.;
RT "Primary structures of neuropeptides isolated from the corpora
RT cardiaca of various cetonid beetle species determined by
RT pulsed-liquid phase sequencing and tandem fast atom bombardment mass

RT spectrometry.";
RL Biol. Chem. Hoppe-Seyler 373:133-142(1992).
CC -!- FUNCTION: THIS HORMONE, RELEASED FROM CELLS IN THE CORPORA
CC CARDIACA AFTER THE BEGINNING OF FLIGHT, CAUSES RELEASE OF
CC DIGLYCERIDES FROM THE FAT BODY AND THEN STIMULATES THE FLIGHT
CC MUSCLES TO USE THESE DIGLYCERIDES AS AN ENERGY SOURCE.
CC -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR PIR; A58641; A58641.
DR InterPro; IPR002047; AKH.
DR PROSITE; PS00256; AKH; 1.
KW Neuropeptide; Amidation; Flight; Pyrrolidone carboxylic acid.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 1022 MW; 867AB775AB544736 CRC64;
Query Match 16.7%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 6 S 6
Db 5 S 5

RESULT 68
AKH_TABAT
ID AKH_TABAT STANDARD; PRT; 8 AA.
AC P14595;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Adipokinetic hormone (AKH) (Dipteran corpora cardiaca factor I)
DE (DCC 1).
OS Tabanus atratus (Horse fly).
OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Tabanomorpha; Tabanidae;
OC Tabanus.
OX NCBI_TaxID=7207;
RN [1]
RP SEQUENCE.
RC TISSUE=Corpora cardiaca;
RX MEDLINE=90046758; PubMed=2813385;
RA Jaffe H., Raina A.K., Riley C.T., Fraser B.A., Nachman R.J.,
RA Vogel V.W., Zhang Y.-S., Hayes D.K.;
RT "Primary structure of two neuropeptide hormones with adipokinetic and
RT hypotrehalosemic activity isolated from the corpora cardiaca of horse
RT flies (Diptera).";
RL Proc. Natl. Acad. Sci. U.S.A. 86:8161-8164(1989).
CC -!- FUNCTION: THIS HORMONE, RELEASED FROM CELLS IN THE CORPORA
CC CARDIACA AFTER THE BEGINNING OF FLIGHT, CAUSES RELEASE OF
CC DIGLYCERIDES FROM THE FAT BODY AND THEN STIMULATES THE FLIGHT
CC MUSCLES TO USE THESE DIGLYCERIDES AS AN ENERGY SOURCE.
CC -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR PIR; A33995; A33995.
DR InterPro; IPR002047; AKH.
DR PROSITE; PS00256; AKH; 1.
KW Neuropeptide; Amidation; Flight; Pyrrolidone carboxylic acid.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 949 MW; 86786771A9D1A736 CRC64;
Query Match 16.7%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 G 1
Db 7 G 7

RESULT 69
ALL2_CARMA

ID AL12_CARMA STANDARD; PRT; 8 AA.
AC P81815;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 12.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Multigene family.
SQ SEQUENCE 8 AA; 913 MW; 672879CDBC569AB7 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
|
DB 7 G 7

RESULT 70

ID AL15_CARMA STANDARD; PRT; 8 AA.
AC P81818;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 15.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Multigene family.
FT MOD RES 8
AMIDATION.
SQ SEQUENCE 8 AA; 811 MW; 922879D5AB47687D CRC64;

Query Match 16.7%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
|
DB 2 G 2

RESULT 71

AL16_CARMA

ID AL16_CARMA STANDARD; PRT; 8 AA.
AC P81819;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 16.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Multigene family.
FT MOD RES 8
AMIDATION.
SQ SEQUENCE 8 AA; 813 MW; 7C286B45AB476878 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
|
DB 1 G 1

RESULT 72

ID AL17_CARMA STANDARD; PRT; 8 AA.
AC P81820;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 17.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Multigene family.
FT MOD RES 8
AMIDATION (POTENTIAL).
SQ SEQUENCE 8 AA; 858 MW; C82879D5AB46D865 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
|
DB 1 S 1

RESULT 73

ID	ALL3_CVDPO	STANDARD;	PRT;	8 AA.
AC	P82154;			
DT	30-MAY-2000	(Rel. 39, Created)		
DT	30-MAY-2000	(Rel. 39, Last sequence update)		
DT	30-MAY-2000	(Rel. 39, Last annotation update)		
DB	Cydiaastatin 3.			
OS	Cydia pomonella	(Codling moth).		
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Lepidoptera; Glossata; Diptera;			
OC	Tortricidae; Tortricidae; Olethreutinae; Cydia.			
OX	NCBI_TaxID=82600;			
RN	[1]			
RP	SEQUENCE.			
RC	TISSUE=Larva;			
RX	MEDLINE=99054539; PubMed=9392829;			
RA	Duve H., Johnson A.H., Maestro J.-L., Scott A.G., Winstanley D.,			
RA	Davey M., East P.D., Thorpe A.;			
RA	"Lepidopteran peptides of the allatostatin superfamily.";			
RL	Peptides 18:1301-1309(1997).			
CC	-1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.			
KW	Neuropeptide; Amidation.			
FT	MOD.RES 8	AMIDATION.		
SEQ	SEQUENCE 8 AA; 926 MW; C82879DSAB477415 CRC64;			
	Query Match	16.7%;	Score 1; DB 1; Length 8;	
	Best Local Similarity	100.0%;	Pred. No. 1.3e+05;	
	Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	6 S 6			
DB	1 S 1			
RESULT 76				
ALL4_CALVO				
ID	ALL4_CALVO	STANDARD;	PRT;	8 AA.
AC	P41840;			
DT	01-NOV-1995	(Rel. 32, Created)		
DT	01-NOV-1995	(Rel. 32, Last sequence update)		
DT	30-MAY-2000	(Rel. 39, Last annotation update)		
DE	Callatostatin 4 (Leu-callatostatin 4).			
OS	Calliphora vomitoria	(Blue blowfly).		
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;			
OC	Calliphoridae; Calliphora.			
OX	NCBI_TaxID=27454;			
RN	[1]			
RP	SEQUENCE.			
RC	TISSUE=Thoracic ganglion;			
RX	MEDLINE=93211890; PubMed=8460157;			
RA	Duve H., Johnson A.H., Scott A.G., Yu C.G., Yagi K.J., Tobe S.S.,			
RA	Thorpe A.;			
RT	"Callatostatins: neuropeptides from the blowfly Calliphora vomitoria			
RT	with sequence homology to cockroach allatostatins.";			
RN	Proc. Natl. Acad. Sci. U.S.A. 90:2456-2460(1993).			
RL	[2]			
RP	CHARACTERIZATION.			
RX	MEDLINE=942931167; PubMed=8020069;			
RA	Duve H., Thorpe A.;			
RT	"Distribution and functional significance of Leu-callatostatins in			
RT	the blowfly Calliphora vomitoria.";			
RL	Cell Tissue Res. 276:367-379(1994).			
CC	-1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR AND PLAY			
CC	A ROLE IN THE INTEGRATION OF INFORMATION WITHIN THE BRAIN. MAY BE			
CC	INVOLVED IN THE CONTROL OF VISCERAL MUSCLES DUE TO ITS ABILITY TO			
CC	BEHAVE AS POTENT INHIBITORS OF PERISTALTIC MOVEMENTS. MAY ALSO			
CC	FULFILL A NEUROHORMONAL ROLE ON MUSCLES OF THE GUT AND HEART.			
CC	-1- TISSUE SPECIFICITY: BRAIN, SUBESOPHAGEAL GANGLION, RETROCEBRAL			
CC	COMPLEX, THORACICO-ABDOMINAL GANGLION, PERIPHERAL NEUROSECRETORY			
CC	SYSTEM AND INTESTINE.			
CC	-1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.			
CC	PIR; D47393; D47393.			
KW	Neuropeptide; Amidation.			

```

FT MOD_RES      8      8      AMIDATION.
FT UNSURE       1      1      OR N.
SQ SEQUENCE     8 AA; 954 MW; D32879D5AB47740A CRC64;

Query Match      16.7%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      6 S 6
      |
DB      5 S 5

RESULT 77
ALL4 CYDPO
ID ALL4 CYDPO STANDARD; PRT; 8 AA.
AC P82155;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cydiaastatin 4.
OS Cydia pomonella (Codling moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Tortricidae; Tortricidae; Olethreutinae; Cydia.
OX NCBI_TaxID=82600;
RN [1]
RP SEQUENCE.
RC TISSUE=Larva;
RX MEDLINE=98054539; PubMed=9392829;
RA Duve H., Johnsen A.H., Maestros J.-L., Scott A.G., Winstanley D.,
RA Davey M., East P.D., Thorpe A.;
RT "Lepidopteran peptides of the allatostatin superfamily.";
RL Peptides 18:1301-1309(1997).
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES      8      8      AMIDATION.
SQ SEQUENCE     8 AA; 910 MW; 922879D5AB47740D CRC64;

Query Match      16.7%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      6 S 6
      |
DB      5 S 5

RESULT 78
ALL5 CALVO
ID ALL5 CALVO STANDARD; PRT; 8 AA.
AC P41841;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Callatostatin 5 [Met-callatostatin 1] ([Hyp3]Met-callatostatin).
OS Calliphora vomitoria (Blue blowfly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
OC Calliphoridae; Calliphora.
OX NCBI_TaxID=27454;
RN [1]
RP SEQUENCE.
RX MEDLINE=93211980; PubMed=8460157;
RA Duve H., Johnsen A.H., Scott A.G., Yu C.G., Yagi K.J., Tobe S.S.,
RA Thorpe A.;
RT "Callatostatins: neuropeptides from the blowfly Calliphora vomitoria
RT with sequence homology to cockroach allatostatins.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:2456-2460(1993).
RN [2]
RP CHARACTERIZATION, AND HYDROXYLATION.
RC TISSUE=Head;
RX MEDLINE=94342269; PubMed=8063725;

```

```

RA Duve H., Johnsen A.H., Scott A.G., East P., Thorpe A.;
RT "[Hyp3]Met-callatostatin. Identification and biological properties of
RT a novel neuropeptide from the blowfly Calliphora vomitoria.";
RL J. Biol. Chem. 269:21059-21066(1994).
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR AND PLAY
CC A ROLE IN THE INTEGRATION OF INFORMATION WITHIN THE BRAIN. MAY BE
CC INVOLVED IN THE CONTROL OF VISCERAL MUSCLES DUE TO ITS ABILITY TO
CC BEHAVE AS POTENT INHIBITORS OF PERISTALTIC MOVEMENTS. MAY ALSO
CC FULFILL A NEUROHORMONAL ROLE ON MUSCLES OF THE GUT AND HEART.
CC -!- TISSUE SPECIFICITY: NEURONS WITHIN BRAIN AND ABDOMINAL GANGLION.
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
DR PIR; E47393; E47393.
KW Neuropeptide; Amidation; Hydroxylation.
FT MOD_RES      3      3      HYDROXYLATION (20%).
FT MOD_RES      8      8      AMIDATION.
SQ SEQUENCE     8 AA; 883 MW; 7D9879CABB477768 CRC64;

Query Match      16.7%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 G 1
      |
DB      1 G 1

RESULT 79
ALL5 CYDPO
ID ALL5 CYDPO STANDARD; PRT; 8 AA.
AC P82155;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cydiaastatin 5.
OS Cydia pomonella (Codling moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Tortricidae; Tortricidae; Olethreutinae; Cydia.
OX NCBI_TaxID=82600;
RN [1]
RP SEQUENCE.
RC TISSUE=Larva;
RX MEDLINE=98054539; PubMed=9392829;
RA Duve H., Johnsen A.H., Maestros J.-L., Scott A.G., Winstanley D.,
RA Davey M., East P.D., Thorpe A.;
RT "Lepidopteran peptides of the allatostatin superfamily.";
RL Peptides 18:1301-1309(1997).
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES      8      8      AMIDATION.
SQ SEQUENCE     8 AA; 898 MW; 922879CABB58640D CRC64;

Query Match      16.7%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 G 1
      |
DB      3 G 3

RESULT 80
ALL6 CYDPO
ID ALL6 CYDPO STANDARD; PRT; 8 AA.
AC P82157;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cydiaastatin 6.
OS Cydia pomonella (Codling moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Tortricidae; Tortricidae; Olethreutinae; Cydia.

```

```

DX NCBI_TaxID=82600;
RN [1]
RP SEQUENCE.
RX TISSUE=Larva;
RR MEDLINE=98054539; PubMed=9392829;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D.,
RT "Lepidopteran peptides of the allatostatin superfamily.";
RL Peptides 18:1301-1309(1997).
CC -|- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 8
SQ SEQUENCE 8 AA; 936 MW; 0B2879C45B573767 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred.No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DY 1 G 1
DB 7 G 7

RESULT 81
ALL7 CARMA STANDARD; PRT; 8 AA.
AC P81809; P81804; P81810;
CT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
JC Carcinustatin 7 [Contains: Carcinustatin 6; Carcinustatin 1].
DS Carcinus maenas (Common shore crab) (Green crab).
DC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
XC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
XC Eubrachyura; Portunoidae; Portunidae; Carcinus.
NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RX TISSUE=Cerebral ganglion, and Thoracic ganglion;
RR MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RT "Isolation and identification of multiple neuropeptides of the
allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -|- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
C -|- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
W Neuropeptide; Amidation; Multigene family.
T PEPTIDE 1 8 CARCINUSTATIN 7.
T PEPTIDE 2 8 CARCINUSTATIN 6.
T PEPTIDE 4 8 CARCINUSTATIN 1.
T MOD_RES 8
FT MOD_RES 8
SQ SEQUENCE 8 AA; 825 MW; 922879CDB4775BD CRC64;

Query Match 16.7%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred.No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DY 6 S 6
DB 2 S 2

RESULT 82
ALL8 CARMA STANDARD; PRT; 8 AA.
AC P81811;
CT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
JC Carcinustatin 8.
DS Carcinus maenas (Common shore crab) (Green crab).
XC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;

```

```

OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RX TISSUE=Cerebral ganglion, and Thoracic ganglion;
RR MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RT "Isolation and identification of multiple neuropeptides of the
allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -|- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
C -|- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
W Neuropeptide; Amidation; Multigene family.
FT MOD_RES 8
SQ SEQUENCE 8 AA; 795 MW; 922879CDB47687D CRC64;

Query Match 16.7%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred.No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 G 1
DB 2 G 2

RESULT 83
ALL9 CARMA STANDARD; PRT; 8 AA.
AC P81812;
CT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
JC Carcinustatin 9.
DS Carcinus maenas (Common shore crab) (Green crab).
XC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RX TISSUE=Cerebral ganglion, and Thoracic ganglion;
RR MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RT "Isolation and identification of multiple neuropeptides of the
allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -|- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
C -|- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
W Neuropeptide; Amidation; Multigene family.
FT MOD_RES 8
SQ SEQUENCE 8 AA; 781 MW; 7C2879CDB476878 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred.No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 G 1
DB 1 G 1

RESULT 84
ANG2 BOTJA STANDARD; PRT; 8 AA.
AC Q10582;
CT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 15-DEC-1998 (Rel. 37, Last annotation update)
JC Angiotensin-like peptide II (Fragment).
XC Carcinus maenas (Common shore crab) (Green crab).
XC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
 OC Viperidae; Crotalinae; Bothrops.
 OX NCBI_TaxID=8724;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Plasma;
 RX MEDLINE=96208932; PubMed=8829801;
 RA Borgheresi R.A.M.B., Dalle Luca J., Carmona E., Picarelli Z.P.;
 RT "Isolation and identification of angiotensin-like peptides from the
 plasma of the snake Bothrops jararaca.";
 RL Comp. Biochem. Physiol. 113B:467-473(1996).
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 DR InterPro: IPR000215; Serpin.
 KW PROSITE: PS00284; SERPIN; PARTIAL.
 FT Vasoconstrictor; Plasma; Serpin.
 FT NON_TER 8
 SQ SEQUENCE 8 AA; 1046 MW; DDD761E04B34D40A CRC64;

Query Match 16.7%; Score 1; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 H 2
 |
 Db 6 H 6

RESULT 85
 CADI_ENTFA
 ID CADI_ENTFA STANDARD; PRT; 8 AA.
 AC P13268;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-FEB-1991 (Rel. 17, Last annotation update)
 DE Sex pheromone CADI.
 OS Enterococcus faecalis (Streptococcus faecalis).
 OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
 OX NCBI_TaxID=1351;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=85051889; PubMed=6437872;
 RA Mori M., Sagakami Y., Narita M., Isogai A., Fujino M., Kitada C.,
 Craig R.A., Clewell D.B., Suzuki A.;
 RT "Isolation and structure of the bacterial sex pheromone, cADI, that
 induces plasmid transfer in Streptococcus faecalis.";
 RL FEBS Lett. 178:97-100(1984).
 CC -!- FUNCTION: CADI IS INVOLVED IN THE CONJUGATIVE TRANSFER OF THE
 HEMOLYSIN PLASMID PAD1.
 KW Pheromone.
 SQ SEQUENCE 8 AA; 819 MW; 047DD732C735B9C7 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
 |
 Db 3 S 3

RESULT 86
 CCKN_MACEU
 ID CCKN_MACEU STANDARD; PRT; 8 AA.
 AC P30369;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Cholecystokinin (CCK).
 GN CCK.
 OS Macropus eugenii (Tamar wallaby), and
 OS Dasyurus viverrinus (Southeastern quoll).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
 OX NCBI_TaxID=9315, 9279;
 RN [1]
 RP SEQUENCE.
 RC SPECIES=M. eugenii, and D. viverrinus;
 RX TISSUE=Brain;
 RX MEDLINE=88234141; PubMed=3375140;
 RA Fan Z.W., Eng J., Shaw G., Yalow R.S.;
 RT "Cholecystokinin octapeptide purified from brains of Australian
 Macropus";
 RL Peptides 9:429-431(1988).
 CC -!- FUNCTION: THIS PEPTIDE HORMONE INDUCES GALL BLADDER CONTRACTION
 AND THE RELEASE OF PANCREATIC ENZYMES IN THE GUT. ITS FUNCTION
 IN THE BRAIN IS NOT CLEAR.
 CC -!- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.
 DR PIR: A43001; A43001.
 DR PIR: PQ0012; PQ0012.
 DR InterPro: IPR001651; Gastrin.
 DR PROSITE: PS00259; GASTRIN; 1.
 KW Amidation; Sulfation; Hormone.
 FT MOD_RES 2 2 SULFATION.
 FT MOD_RES 8 8 AMIDATION.
 SQ SEQUENCE 8 AA; 1064 MW; DDCAA68378768B5A CRC64;

Query Match 16.7%; Score 1; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 |
 Db 4 G 4

RESULT 87
 CLP_THICU
 ID CLP_THICU STANDARD; PRT; 8 AA.
 AC P80488;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Chemolithotroph-specific protein (Fragment).
 OS Thiobacillus cuprinus.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Comamonadaceae; Thiomonas.
 OX NCBI_TaxID=36860;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=DSM 5494;
 RA Marin I., Amato A.M., Jerez C.A., Abad J.P.;
 RL Submitted (SEP-1995) to the SWISS-PROT data bank.
 CC -!- MISCELLANEOUS: FOUND SPECIFICALLY IN CELLS CULTURED
 CC -!- CHEMOLITHOTROPHICALLY.
 FT NON_TER 8
 SQ SEQUENCE 8 AA; 785 MW; 91487B06DDC2D76D CRC64;

Query Match 16.7%; Score 1; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 |
 Db 7 G 7

RESULT 88
 COW2_CONPU
 ID COW2_CONPU STANDARD; PRT; 8 AA.
 AC P58785;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Leu-contryphan-P.
 OS Conus purpurascens (Purple cone).

OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
 OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
 OC Neogastropoda; Conoidea; Conidae; Conus.
 DX NCBI_TaxID=41690;

[1]

SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.

STRAIN=Clipperton Island; TISSUE=Venom;

MEDLINE=9938839; PubMed=10461743;

Jacobson R.B., Jimenez E.C., De la Cruz R.G.C., Gray W.R., Cruz L.J.,

Olivera B.M.;

"A novel D-leucine-containing Conus peptide: diverse conformational

dynamics in the contryphan family.";

J. Pept. Res. 54:93-99(1999).

-1- SUBCELLULAR LOCATION: Secreted.

-1- TISSUE SPECIFICITY: Expressed by the venom duct.

-1- MASS SPECTROMETRY: MW=888.4; METHOD=LSIMS.

-1- SIMILARITY: BELONGS TO THE CONTRYPHAN FAMILY.

Toxin; Hydroxylation; D-amino acid.

DISULFID 2

MOD RES 4 4

SEQUENCE 8 AA; 830 MW; 75A367672732CEB8 CRC64;

D-LEUCINE.

Query Match 16.7%; Score 1; DB 1; Length 8;

Best Local Similarity 100.0%; Pred. No. 1.3e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

1 G 1

RESULT 90

CPDI_ENTFA

ID CPDI_ENTFA

STANDARD;

PRT;

8 AA.

AC P13269;

DT 01-JAN-1990 (Rel. 13, Created)

DT 01-JAN-1990 (Rel. 13, Last sequence update)

DT 01-FEB-1991 (Rel. 17, Last annotation update)

DE Sex pheromone CPDI.

OS Enterococcus faecalis (Streptococcus faecalis).

OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.

OX NCBI_TaxID=1351;

RN [1]

RP SEQUENCE.

RX MEDLINE=65040388; PubMed=6436978;

RA Suzuki A., Mori M., Sagakami Y., Isogai A., Fujino M., Kitada C.,

RA Craig R.A., Clewell D.B.;

RT "Isolation and structure of bacterial sex pheromone, cPDI.";

RL Science 226:849-850(1984).

CC -1- FUNCTION: CPDI IS INVOLVED IN THE CONJUGATIVE TRANSFER OF THE

CC BACTERIOCIN PLASMID PPDI.

KW Pheromone.

SQ SEQUENCE 8 AA; 913 MW; 8655B729C682C729 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 8;

Best Local Similarity 100.0%; Pred. No. 1.3e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

6 S 6

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

RESULT 91

FARI_PANRE

ID FARI_PANRE

STANDARD;

PRT;

8 AA.

AC P41872;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 01-NOV-1995 (Rel. 32, Last annotation update)

DE FMRamide-like neurotoxin PFI (SDPNFLR-amide).

OS Panagrellus redivivus.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;

OC Panagrolaimoidea; Panagrolaimidae; Panagrellus.

OX NCBI_TaxID=6233;

RN [1]

RP SEQUENCE.

RX MEDLINE=93027659; PubMed=1408999;

RA Geary T.G., Price D.A., Bowman J.W., Winterrowd C.A., Mackenzie C.D.,

RA Garrison R.D., Williams J.F., Friedman A.R.;

RT "Two FMRamide-like peptides from the free-living nematode

Panagrellus redivivus.";

RL Peptides 13:209-214(1992).

CC -1- FUNCTION: MYOACTIVE.

CC -1- TISSUE SPECIFICITY: NERVE CORDS AND PAIRED GROUPS OF CELLS LOCATED

CC CAUDALLY TO THE BASE OF THE PHARYNX.

CC -1- SIMILARITY: BELONGS TO THE FARI (FMRAMIDE RELATED PEPTIDE)

CC FAMILY.

KW Neurotoxin; Amidation.

FT MOD RES 8

SQ SEQUENCE 8 AA; 995 MW; C6D40729C4576AB5 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 8;

Best Local Similarity 100.0%; Pred. No. 1.3e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

6 S 6

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

7 S 7

RESULT 92

FAR1_PENMO
ID FAR1_PENMO STANDARD; PRT; 8 AA.
AC P8316;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE FMRFamide-like neuropeptide FLPI (GDRNFLAF-amide).
OS Peneaus monodon (Peneid shrimp).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;
OC Penaeidae; Peneaus.
OX NCBI_TaxID=6687;
RN [1];
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE=Eye stalk;
RX MEDLINE=21956277; PubMed=11959015;
RA Sithigorngul P., Pupum J., Krungkarn C., Longyant S.,
RA Chaivuthangkura P., Sithigorngul W., Petsom A.;
RT "Seven novel FMRFamide-like neuropeptide sequences from the eyestalk
of the giant tiger prawn *Peneaus monodon*.";
RL Comp. Biochem. Physiol. 131B:325-337(2002).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MASS SPECTROMETRY: MW=1024.8; METHOD=MALDI.
CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 1024 MW; 72D40729C4540AAB CRC64;

Query Match 16.7%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
DB 1 G 1

RESULT 93
FAR3_HOMAM
ID FAR3_HOMAM STANDARD; PRT; 8 AA.
AC P41486;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE FMRFamide-like neuropeptide 3 (PLI 3) (F2).
OS Homarus americanus (American lobster).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
OC Nephropoidea; Nephropidae; Homarus.
OX NCBI_TaxID=6706;
RN [1];
RP SEQUENCE.
RC TISSUE=Pericardial organs;
RX MEDLINE=88116164; PubMed=3429714;
RA Timmer B.A., Kobierski L.A., Kravitz E.A.;
RT "Purification and characterization of FMRFamide-like immunoreactive
substances from the lobster nervous system: Isolation and sequence
analysis of two closely related peptides.";
RL J. Comp. Neurol. 266:16-26(1987).
CC -1- MISCELLANEOUS: PERICARDIAL ORGANS RELEASE THIS PEPTIDE WITH 100 MW
POTASSIUM IN THE PRESENCE OF CALCIUM.
CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 1054 MW; C6D40729C4540AAB5 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6

Db 1 S 1

RESULT 94
FAR7_ASCSU
ID FAR7_ASCSU STANDARD; PRT; 8 AA.
AC P43171;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE FMRFamide-like neuropeptide AF7.
OS Ascaris suum (Pig roundworm) (Ascaris lumbricoides).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC Ascarididae; Ascaris.
OX NCBI_TaxID=6253;
RN [1];
RP SEQUENCE.
RX MEDLINE=95380362; PubMed=7651904;
RA Cowden C., Stretton A.O.W.;
RT "Eight novel FMRFamide-like neuropeptides isolated from the nematode
Ascaris suum.";
RL Peptides 16:491-500(1995).
CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 963 MW; 9CD40059D417687D CRC64;

Query Match 16.7%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
DB 2 G 2

RESULT 95
FAR8_CALVO
ID FAR8_CALVO STANDARD; PRT; 8 AA.
AC P41833;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE CalliprFamide 8.
OS Calliphora vomitoria (Blue blowfly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
OC Calliphoridae; Calliphora.
OX NCBI_TaxID=27454;
RN [1];
RP SEQUENCE.
RC TISSUE=Thoracic ganglion;
RX MEDLINE=92196111; PubMed=1549595;
RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,
RA Rehfeld J.F., Thorpe A.;
RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
neuropeptides (designated calliprFamides) from the blowfly
Calliphora vomitoria.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
FAMILY.
KW PIR; H41978; H41978.
DR Neuropeptide; Amidation.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 957 MW; 72D4069CAA44DD8 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1

```

Db      1 G 1
RESULT 96
FUSS_FUSSO
ID_FUSS_FUSSO STANDARD; PRT; 8 AA.
AC P01010-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Allergen Fus s I3596* (Fragment).
OS Fusarium solani (subsp. Pisi) (Nectria haematococcal).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Nectriaceae; Nectria.
OX NCBI_TaxID=70791;
[1]
SEQUENCE.
RC STRAIN=IARI 3596; TISSUE=Mycelium;
RA Verma J., Gangal S.V.;
RL Submitted (JUL-1997) to the SWISS-PROT data bank.
KW Allergen.
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 898 MW; C372C441F5B690A1 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 H 2
Db      5 H 5

RESULT 97
HTF1_PERAM
ID HTF1_PERAM STANDARD; PRT; 8 AA.
AC P04548;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypertrehalosaemic factor I (Neuropeptide M-I) (Periplanetin CC-I)
DE (Pea-CAH-I) (Leb-CC-I) (Hypertrehalosaemic neuropeptide I).
OS Periplaneta americana (American cockroach),
OS Leptinotarsa decemlineata (Colorado potato beetle), and
OS Blattella orientalis (Oriental cockroach).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blattodea;
OC Blattidae; Periplaneta.
OX NCBI_TaxID=6978, 7539, 6976;
[1]
SEQUENCE.
RC SPECIES=P.americana;
RC MEDLINE=85046530; PubMed=6548628;
RA Witten J.L., Schaffer M.H., O'Shea M., Cook J.C., Hemling M.E.,
RA Rinehart K.L., Jr.;
DT "Structures of two cockroach neuropeptides assigned by fast atom
RT bombardment mass spectrometry."
RL Biochem. Biophys. Res. Commun. 124:350-358(1984).
[2]
SEQUENCE.
RC SPECIES=P.americana;
RC MEDLINE=84296179; PubMed=6591205;
RA Scarborough R.M., Jamieson G.C., Kalish F., Kramer S.J., McEnroe G.A.,
RA Miller C.A., Schooley D.A.;
RT "Isolation and primary structure of two peptides with
RT cardioacceleratory and hyperglycemic activity from the corpora
RT cardiaca of Periplaneta americana."
RL Proc. Natl. Acad. Sci. U.S.A. 81:5575-5579(1984).
[3]
SEQUENCE.
RC SPECIES=L.decemlineata; TISSUE=Corpora cardiaca;
RC MEDLINE=90160053; PubMed=2576128;

```

```

RA Gaede G., Kellner R.;
RT "The metabolic neuropeptides of the corpus cardiacum from the potato
RL beetle and the American cockroach are identical."
RN Peptides 10:1287-1289(1989).
[4]
SEQUENCE.
RC SPECIES=B.orientalis; TISSUE=Corpora cardiaca;
RC MEDLINE=90253659; PubMed=2340112;
RA Gaede G., Rinehart K.L., Jr.;
RT "Primary structures of hypertrehalosaemic neuropeptides isolated from
RT the corpora cardiaca of the cockroaches Leucophaea maderae,
RT Gromphadorhina portentosa, Blattella germanica and Blatta orientalis
RT and of the stick insect Extatosoma tiaratum assigned by tandem fast
RT atom bombardment mass spectrometry."
RL Biol. Chem. Hoppe-Seyler 371:345-354(1990).
CC -!- FUNCTION: HYPERTREHALOSAEMIC FACTORS ARE NEUROPEPTIDES THAT
CC ELEVATE THE LEVEL OF TREHALOSE IN THE HEMOLYMPH (TREHALOSE IS
CC THE MAJOR CARBOHYDRATE IN THE HEMOLYMPH OF INSECTS).
CC -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR PIR; A44960; A44960.
DR PIR; A49823; A49823.
DR PIR; S08995; S08995.
DR InterPro; IPR002047; AKH.
DR PROSITE; PS00256; AKH; 1.
KW Neuropeptide; Amidation; Pyrrolidone carboxylic acid.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 991 MW; 86745775B9C452D6 CRC64;

Query Match 16.7%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      6 S 6
Db      5 S 5

RESULT 98
HTF_TENNO
ID HTF_TENNO STANDARD; PRT; 8 AA.
AC P25419;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypertrehalosaemic factor (HOTH) (Hypertrehalosemic neuropeptide).
OS Tenebrio molitor (Yellow mealworm), and
OS Zophobas rugipes.
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
OC Tenebrionidae; Tenebrio.
OX NCBI_TaxID=7067, 7075;
[1]
SEQUENCE.
RC SPECIES=T.molitor, and Z.rugipes;
RC TISSUE=Corpora cardiaca;
RX MEDLINE=90341081; PubMed=2381871;
RA Gaede G., Rosinski G.;
RT "The primary structure of the hypertrehalosemic neuropeptide from
RT tenebrionid beetles: a novel member of the AKH/RPCH family."
RL Peptides 11:455-459(1990).
CC -!- FUNCTION: HYPERTREHALOSAEMIC FACTORS ARE NEUROPEPTIDES THAT
CC ELEVATE THE LEVEL OF TREHALOSE IN THE HEMOLYMPH (TREHALOSE IS
CC THE MAJOR CARBOHYDRATE IN THE HEMOLYMPH OF INSECTS).
CC -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR PIR; A43976; A43976.
DR PIR; B43976; B43976.
DR InterPro; IPR002047; AKH.
DR PROSITE; PS00256; AKH; 1.
KW Neuropeptide; Amidation; Pyrrolidone carboxylic acid.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 1005 MW; 86745775B9C44736 CRC64;

```

Query Match 16.7%; Score 1; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred.No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
 |
 Db 5 S 5

RESULT 99
 LCK1 LEUMA
 ID LCK1 LEUMA STANDARD; PRT; 8 AA.
 AC P21140;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 01-MAY-1991 (Rel. 18, Last annotation update)
 DE Leucokinin I (L-I).
 OS Leucophaea maderae (Madeira cockroach).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;
 OC Blaberidae; Leucophaea.
 OX NCBI_TaxID=6988;
 RN [1]
 RP SEQUENCE, AND SYNTHESIS.
 RC TISSUE=Head;
 RA Holman G.M., Cook B.J., Nachman R.J.;
 RT "Isolation, primary structure and synthesis of two neuropeptides
 from Leucophaea maderae: members of a new family of
 Cephalomyotropins.";
 RL Comp. Biochem. Physiol. 84C:205-211(1986).
 CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
 CC -!- ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).
 CC -!- SIMILARITY: TO THE OTHER LEUCOKININS.
 KW Neuropeptide; Amidation.
 FT MOD_RES 8
 SQ SEQUENCE 8 AA; 893 MW; DC6365B449CDC76A CRC64;

Query Match 16.7%; Score 1; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred.No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
 |
 Db 6 S 6

RESULT 100
 LCK2 LEUMA
 ID LCK2 LEUMA STANDARD; PRT; 8 AA.
 AC P21141;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 01-MAY-1991 (Rel. 18, Last annotation update)
 DE Leucokinin II (L-II).
 OS Leucophaea maderae (Madeira cockroach).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;
 OC Blaberidae; Leucophaea.
 OX NCBI_TaxID=6988;
 RN [1]
 RP SEQUENCE, AND SYNTHESIS.
 RC TISSUE=Head;
 RA Holman G.M., Cook B.J., Nachman R.J.;
 RT "Isolation, primary structure and synthesis of two neuropeptides
 from Leucophaea maderae: members of a new family of
 Cephalomyotropins.";
 RL Comp. Biochem. Physiol. 84C:205-211(1986).
 CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
 CC -!- ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).
 CC -!- SIMILARITY: TO THE OTHER LEUCOKININS.
 KW Neuropeptide; Amidation.
 FT MOD_RES 8
 AMIDATION.

SQ SEQUENCE 8 AA; 852 MW; DC6365A5B9C676A CRC64;

Query Match 16.7%; Score 1; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred.No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 |
 Db 3 G 3

Search completed: October 21, 2003, 18:24:31
 Job time : 13.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:19:26 ; Search time 47.6667 Seconds

(without alignments)
32.482 Million cell updates/sec

Title: US-10-057-890A-15

Perfect score: 6

Sequence: 1 GHHHS 6

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 830525 segs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1349

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Listing first 150 summaries

Database :

SPTREMBL 23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_ivirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3	50.0	8	4	Q15894 homo sapien
2	3	33.3	7	10	P82445 nicotiana t
3	2	33.3	7	12	Q65578 bovine herp
4	2	33.3	8	12	Q90345 hepatis g
5	2	33.3	9	2	Q7410 escherichia
6	2	33.3	9	4	Q9H4M8 homo sapien
7	2	33.3	9	4	Q9UC36 homo sapien
8	2	33.3	9	8	P92072 euhadra her
9	2	33.3	9	11	Q924N8 mus musculu
10	2	33.3	9	11	Q08979 mus musculu
11	2	33.3	9	11	Q8CG39 rattus norv
12	2	33.3	10	2	Q9UN49 staphylococ
13	2	33.3	10	2	Q47561 escherichia
14	2	33.3	10	3	Q9UW2 schizophyll
15	2	33.3	10	3	Q9HDS2 aspergillus
16	2	33.3	10	5	P82223 bombyx mori

17	2	33.3	10	5	P82224	P82224 bombyx mori
18	2	33.3	10	10	P81899	P81899 prunus dulc
19	2	33.3	10	15	Q85563	Q85563 moloney mur
20	2	16.7	5	10	Q99007	Q99007 hordeum vul
21	1	16.7	6	10	P82181	P82181 spinacia ol
22	1	16.7	6	10	P82541	P82541 spinacia ol
23	1	16.7	6	10	P82182	P82182 spinacia ol
24	1	16.7	7	2	Q8KMS3	Q8KMS3 klebsiella
25	1	16.7	7	2	Q47505	Q47505 escherichia
26	1	16.7	7	2	P70804	P70804 azotobacter
27	1	16.7	7	2	Q47029	Q47029 enterobacte
28	1	16.7	7	2	P72081	P72081 nocardia la
29	1	16.7	7	2	Q07354	Q07354 synecococc
30	1	16.7	7	6	Q28742	Q28742 oryctolagus
31	1	16.7	7	8	Q9Y182	Q9Y182 gnatholebia
32	1	16.7	7	8	Q8MFY6	Q8MFY6 taraxacum (
33	1	16.7	7	9	Q95945	Q95945 saccharomyc
34	1	16.7	7	10	P93233	P93233 lycopersico
35	1	16.7	7	10	Q49223	Q49223 glycine max
36	1	16.7	7	10	Q9C5B3	Q9C5B3 arabidopsis
37	1	16.7	7	11	Q8K3H6	Q8K3H6 rattus norv
38	1	16.7	7	11	Q63480	Q63480 rattus norv
39	1	16.7	7	11	O55184	O55184 rattus norv
40	1	16.7	7	11	Q63668	Q63668 rattus norv
41	1	16.7	7	12	Q66205	Q66205 transmissib
42	1	16.7	7	13	Q8JJ20	Q8JJ20 gallus gall
43	1	16.7	7	13	Q42564	Q42564 fugu rubrip
44	1	16.7	7	15	Q07624	Q07624 rous sarcom
45	1	16.7	8	2	Q45615	Q45615 bacillus su
46	1	16.7	8	2	Q9AGP4	Q9AGP4 arthrobacte
47	1	16.7	8	2	Q9R3X0	Q9R3X0 planktothri
48	1	16.7	8	2	Q9R5R2	Q9R5R2 shigella dy
49	1	16.7	8	2	Q64885	Q64885 klebsiella
50	1	16.7	8	2	Q93SR0	Q93SR0 staphylococ
51	1	16.7	8	2	Q8RJ10	Q8RJ10 streptomyce
52	1	16.7	8	2	Q44463	Q44463 rhizobiales
53	1	16.7	8	2	P77556	P77556 escherichia
54	1	16.7	8	2	Q9ZE29	Q9ZE29 buchnera ap
55	1	16.7	8	2	O09258	O09258 synecococc
56	1	16.7	8	2	Q8RSR3	Q8RSR3 lactobacill
57	1	16.7	8	2	Q8X3K1	Q8X3K1 prochloroco
58	1	16.7	8	2	O87471	O87471 haemophilus
59	1	16.7	8	2	Q56140	Q56140 streptococc
60	1	16.7	8	2	Q56246	Q56246 thermophili
61	1	16.7	8	2	Q56429	Q56429 thermus the
62	1	16.7	8	2	Q9R9S0	Q9R9S0 bacillus su
63	1	16.7	8	2	Q93822	Q93822 pseudomonas
64	1	16.7	8	2	Q45889	Q45889 clostridium
65	1	16.7	8	2	Q9R5R0	Q9R5R0 shigella dy
66	1	16.7	8	2	Q9R9C2	Q9R9C2 borrelia bu
67	1	16.7	8	2	Q47273	Q47273 escherichia
68	1	16.7	8	2	Q9R4M3	Q9R4M3 enterococcu
69	1	16.7	8	2	Q93454	Q93454 thiobacillu
70	1	16.7	8	2	P72221	P72221 pseudomonas
71	1	16.7	8	2	O85406	O85406 coxiella bu
72	1	16.7	8	2	Q9R5L7	Q9R5L7 clostridium
73	1	16.7	8	2	P83158	P83158 anabaena sp
74	1	16.7	8	3	Q05403	Q05403 saccharomyc
75	1	16.7	8	3	O13591	O13591 saccharomyc
76	1	16.7	8	3	P87225	P87225 saccharomyc
77	1	16.7	8	3	Q9HDS4	Q9HDS4 aspergillus
78	1	16.7	8	3	Q9URB9	Q9URB9 saccharomyc
79	1	16.7	8	3	P82858	P82858 puccinia re
80	1	16.7	8	4	Q9BYV5	Q9BYV5 homo sapien
81	1	16.7	8	4	Q15889	Q15889 homo sapien
82	1	16.7	8	4	Q15900	Q15900 homo sapien
83	1	16.7	8	4	O16428	O16428 homo sapien
84	1	16.7	8	4	Q9UHK1	Q9UHK1 homo sapien
85	1	16.7	8	4	Q15895	Q15895 homo sapien
86	1	16.7	8	4	Q9Y4X6	Q9Y4X6 homo sapien
87	1	16.7	8	4	Q15901	Q15901 homo sapien
88	1	16.7	8	4	Q9Y4J4	Q9Y4J4 homo sapien
89	1	16.7	8	4	Q15890	Q15890 homo sapien

90	1	16.7	8	4	Q9UDZ4
91	1	16.7	8	4	Q15898
92	1	16.7	8	4	Q15893
93	1	16.7	8	4	Q15888
94	1	16.7	8	4	Q16468
95	1	16.7	8	4	Q9UCN4
96	1	16.7	8	4	Q9UMC7
97	1	16.7	8	4	Q9HCQ0
98	1	16.7	8	4	Q9P285
99	1	16.7	8	4	Q8TF70
100	1	16.7	8	4	Q9P0K3
101	1	16.7	8	4	Q9UMH9
102	1	16.7	8	4	Q60773
103	1	16.7	8	4	Q8IUH8
104	1	16.7	8	5	Q9TWH6
105	1	16.7	8	5	Q94623
106	1	16.7	8	5	Q8MUN6
107	1	16.7	8	5	Q02032
108	1	16.7	8	5	Q94695
109	1	16.7	8	5	P82685
110	1	16.7	8	5	P82686
111	1	16.7	8	5	P82687
112	1	16.7	8	5	P82688
113	1	16.7	8	5	P82689
114	1	16.7	8	6	Q9TT78
115	1	16.7	8	6	Q9BFA7
116	1	16.7	8	6	Q9BF82
117	1	16.7	8	6	Q9BFC2
118	1	16.7	8	6	Q9BF90
119	1	16.7	8	6	Q9BF81
120	1	16.7	8	6	Q18854
121	1	16.7	8	6	Q9BFA1
122	1	16.7	8	6	Q9BF87
123	1	16.7	8	6	Q9BF89
124	1	16.7	8	6	Q9BFA0
125	1	16.7	8	6	Q9BFA8
126	1	16.7	8	6	Q9XSY1
127	1	16.7	8	6	Q9BFA9
128	1	16.7	8	6	Q9BFB2
129	1	16.7	8	6	Q9BFB5
130	1	16.7	8	6	Q9BFB6
131	1	16.7	8	6	Q9BFB3
132	1	16.7	8	6	Q9BF88
133	1	16.7	8	6	Q9BF95
134	1	16.7	8	6	Q9BF99
135	1	16.7	8	6	Q9MYL5
136	1	16.7	8	6	Q9BFB4
137	1	16.7	8	6	Q9BFC3
138	1	16.7	8	6	Q9BFA4
139	1	16.7	8	6	Q9GMH3
140	1	16.7	8	6	Q9BFA2
141	1	16.7	8	6	Q9BF85
142	1	16.7	8	6	Q9BFC1
143	1	16.7	8	6	Q9BFB9
144	1	16.7	8	6	Q9BF96
145	1	16.7	8	6	Q9BFB0
146	1	16.7	8	6	Q9BF86
147	1	16.7	8	6	Q02831
148	1	16.7	8	6	Q9BFB7
149	1	16.7	8	6	Q9BFB9
150	1	16.7	8	6	Q9BFB1

ALIGNMENTS

ESULT 1
15894
D Q15894 PRELIMINARY; PRT; 8 AA.
C Q15894; (TREMBLrel. 01, Created)
T 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
T 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE (Clone XP887B) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Lee C.-C., Yazdani A., Wennert M., Bailey J., Couch L., Xiong M.,
RA Coolbaugh M.I., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,
RA Caskey C.T.H.;
RT "Isolation of chromosome-specific genes by reciprocal probing of
RL Hum. Mol. Genet. 0:0-0(1995).
DR EMBL; L32074; AAA73884.1; -;
FT NON TER 1 8
FT NON TER 8
SQ SEQUENCE 8 AA; 952 MW; EBC735B1EIF1B6D6 CRC64;

Query Match 50.0%; Score 3; DB 4; Length 8;
Best Local Similarity 100.0%; Pred.No.8.3e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 HHS 6
Db 4 HHS 6

RESULT 2
P82445 PRELIMINARY; PRT; 7 AA.
AC P82445;
DT 01-JUN-2000 (TREMBLrel. 14, Created)
DT 01-JUN-2000 (TREMBLrel. 14, Last sequence update)
DE 10 kDa cell wall protein (Fragment).
OC Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
OC Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE.
RC STRAIN=cv. PETIT HAVANA;
RA Blee K.A., Bonham V.A., Mitchell G.P., Robertson D., Slabas A.R.,
RA Wojtaszek P., Bolwell G.P.;
RT "Proteomic study of secondary cell wall proteins from transformed
RT tobacco culture";
RL Planta 0:0-0(2000).
CC -I- SUBCELLULAR LOCATION: CELL WALL.
CC -I- TISSUE SPECIFICITY: XYLEM.
KW Cell wall.
FT NON TER 7
SQ SEQUENCE 7 AA; 758 MW; 69D2C1E862D1B2A0 CRC64;

Query Match 33.3%; Score 2; DB 10; Length 7;
Best Local Similarity 100.0%; Pred.No.8.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GH 2
Db 4 GH 5

RESULT 3
Q65578 PRELIMINARY; PRT; 7 AA.
ID Q65578
AC Q65578;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical 0.8 kDa protein (Fragment).
CS Bovine herpesvirus 1.

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.

OX NCBI_TaxID=10320;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Cooper;

RX MEDLINE=95313343; PubMed=7793062;

RA Vleck C., Benes V., Lu Z., Kutish G.F., Paces V., Rock D.,

RA Letcworth G.J., Schwytzer M.;

RT "Nucleotide sequence analysis of a 30-kb region of the bovine

RT herpesvirus 1 genome which exhibits a colinear gene arrangement with

RL the UL21 to UL4 genes of herpes simplex virus.";

RL Virol. 210:1100-1108 (1995).

DR EMBL: 248053; CAA88130.1; -.

KW Hypothetical protein.

FT NON_TER 1

SQ SEQUENCE 7 AA; 758 MW; 6DD33455B1F1B1C0 CRC64;

Query Match 33.3%; Score 2; DB 12; Length 7;

Best Local Similarity 100.0%; Pred. No. 8.3e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 HS 6

DB 3 HS 4

RESULT 4

Q90345

ID O90345;

AC O90345; PRELIMINARY; PRT; 8 AA.

DT 01-NOV-1998 (TREMBlrel. 08, Created)

DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE Polyprotein (Fragment).

OS Hepatitis G virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OC GBV-C/HGV group.

OX NCBI_TaxID=45255;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SG2813;

RA MEDLINE=99266893; PubMed=1033862;

RA Wong S.B.J., Chan S.H., Ren E.C.;

RT "Diversity of GB virus C/hepatitis G virus isolates in Singapore:

RT predominance of group 2a and the Asian group 3 variant.";

J. Med. Virol. 58:145-153 (1999).

DR EMBL: AF078050; AAC32360.1; -.

FT NON_TER 8

SQ SEQUENCE 8 AA; 799 MW; ECA76DC5B1F2CDD6 CRC64;

Query Match

Best Local Similarity 100.0%; Pred. No. 8.3e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 HS 6

DB 4 HS 5

RESULT 5

Q47410

ID Q47410;

AC Q47410; PRELIMINARY; PRT; 9 AA.

DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE Pot. repA peptide (Fragment).

OS Escherichia coli.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=plasmid NR1;

RX MEDLINE=85160860; PubMed=2580099;

RA Womble D.D., Sampathkumar P., Easton A.M., Luckow V.A., Rownd R.H.;

RT "Transcription of the replication control region of the IncPII R-

RT plasmid NR1 in vitro and in vivo.";

RL J. Mol. Biol. 181:395-410 (1985).

DR EMBL: X02302; CAA26166.1; -.

KW Plasmid.

FT NON_TER 9

SQ SEQUENCE 9 AA; 1055 MW; DCF6A6412CDD1E87D CRC64;

Query Match 33.3%; Score 2; DB 2; Length 9;

Best Local Similarity 100.0%; Pred. No. 8.3e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GH 2

DB 3 GH 4

RESULT 6

Q9H4M8

ID Q9H4M8

AC Q9H4M8; PRELIMINARY; PRT; 9 AA.

DT 01-MAR-2001 (TREMBlrel. 16, Created)

DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)

DE PAR2 (Fragment).

GN NR112.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Peripheral blood;

RA Pentecost B.T., Ling G.;

RT "The human pregnane X receptor promoter complex provides

RT transcriptional starts for a number of PXR related transcripts.";

RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AY007189; AAG23345.1; -.

FT NON_TER 9

SQ SEQUENCE 9 AA; 1129 MW; 82F8E1F1B411B2D1 CRC64;

Query Match

Best Local Similarity 100.0%; Pred. No. 8.3e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HH 3

DB 7 HH 8

RESULT 7

Q9UC36

ID Q9UC36

AC Q9UC36; PRELIMINARY; PRT; 9 AA.

DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE 28 kDa heat shock protein homolog fragment 1 (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE.

RX MEDLINE=92218434; PubMed=1560006;

RA Kato K., Shinohara H., Goto S., Inaguma Y., Morishita R., Asano T.;

RT "Copurification of small heat shock protein with alpha B crystallin

RT from human skeletal muscle.";

```

J. Biol. Chem. 267:7718-7725(1992).
FT NON_TER 1 1
FT NON_TER 9 9
SQ SEQUENCE 9 AA; 1220 MW; 26933415B1F77B43 CRC64;
  Query Match 33.3%; Score 2; DB 4; Length 9;
  Best Local Similarity 100.0%; Pred.No. 8.3e+05;
  Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 HS 6
  ||
DB 5 HS 6

RESULT 8
ID P92072 PRELIMINARY; PRT; 9 AA.
AC P92072;
DT 01-MAY-1997 (TRENBLrel. 03, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE ATPase subunit 8 (Fragment).
XS Euhadra herklotsi.
XS Mitochondrion.
XC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylomatophora;
CC Sigmyrethra; Helicoidea; Bradybaenidae; Euhadra.
CX NCBI_TaxID=58912;
CX [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Hepatopancreas;
RC Yamazaki N., Ueshima R., Terrett J.A., Yokobori S., Kaifu M.,
RX Segawa R., Kobayashi T., Numachi K., Ueda T., Nishikawa K.,
RA Watanabe K., Thomas R.H.;
RT "Evolution of pulmonate gastropod mitochondrial genomes: comparisons
RT of complete gene organization of Euhadra, Cepaea and Albinaria and
RT implications of unusual tRNA secondary structures.";
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
JR EMBL; 271697; CAA96373.1; -.
XT NON_TER 9
Q SEQUENCE 9 AA; 1009 MW; 380CB1F775B736C7 CRC64;

  Query Match 33.3%; Score 2; DB 8; Length 9;
  Best Local Similarity 100.0%; Pred.No. 8.3e+05;
  Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 5 HS 6
  ||
b 7 HS 8

RESULT 9
ID Q924N8 PRELIMINARY; PRT; 9 AA.
AC Q924N8;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Niemann Pick type C1 protein (Fragment).
XS NPC1.
XS Mus musculus (Mouse).
XC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
CX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BLKS;
RC Gevry N.Y., Lacroix D.A., Murphy B.D.;
RT "Niemann-Pick C1 protein gene, partial cds and promoter region.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
JR EMBL; AF184964; AAK83683.1; -.
XT NON_TER 9
Q SEQUENCE 9 AA; 890 MW; 2C4E2DC761E1EDD8 CRC64;

```

```

  Query Match 33.3%; Score 2; DB 11; Length 9;
  Best Local Similarity 100.0%; Pred.No. 8.3e+05;
  Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HS 3
  ||
DB 4 HS 5

RESULT 10
ID O08979 PRELIMINARY; PRT; 9 AA.
AC O08979;
DT 01-JUL-1997 (TRENBLrel. 04, Created)
DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE AML1 protein (Fragment).
XS AML1.
XS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NMRI; TISSUE=Tumor;
RX MEDLINE=97332339; PubMed=9188573;
RA Antoft H.W., Sorensen A.B., Barell C., Schmidt J., Luz A.,
RA Pedersen F.S.;
RT "Stability of AML1 (core) site enhancer mutations in T-lymphomas
RT induced by attenuated SL3-3 murine leukemia virus mutants.";
RL J. Virol. 71:5080-5087(1997).
DR EMBL; Y11802; CAA72496.1; -.
FT NON_TER 1
FT NON_TER 9
SQ SEQUENCE 9 AA; 981 MW; 293E01E865A776D8 CRC64;

  Query Match 33.3%; Score 2; DB 11; Length 9;
  Best Local Similarity 100.0%; Pred.No. 8.3e+05;
  Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GH 2
  ||
DB 6 GH 7

RESULT 11
ID O8CG39 PRELIMINARY; PRT; 9 AA.
AC O8CG39;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Histamine N-tele-methyltransferase (Fragment).
XS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=96342418; PubMed=8750786;
RA Takemura M., Yamauchi K., Yamatodani A.;
RT "Structural analysis of histamine N-methyltransferase gene.";
RL Methods Find. Exp. Clin. Pharmacol. 17:1-4(1995).
DR EMBL; S82579; AAN86745.1; -.
KW Transferase; Methyltransferase.
FT NON_TER 1
SQ SEQUENCE 9 AA; 1060 MW; 4223B1E042CEB053 CRC64;

  Query Match 33.3%; Score 2; DB 11; Length 9;
  Best Local Similarity 100.0%; Pred.No. 8.3e+05;
  Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      5 HS 6
Db      7 HS 8

RESULT 12
QJUN49
ID QJUN49 PRELIMINARY; PRT; 10 AA.
AC QJUN49;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Plasmid pE194 ermC leader region (Fragment).
OS Staphylococcus aureus.
OG Plasmid pE194.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86041917; PubMed=3903662;
RA Narayanan C.S., Dubnau D.;
RT "Evidence for the translational attenuation model: ribosome-binding
RT studies and structural analysis with an in vitro run-off transcript
RT of ermC.";
RL Nucleic Acids Res. 13:7307-7326(1985).
DR EMBL; X03097; CAA26883.1; -.
KW Plasmid.
FT NON_TER
SQ SEQUENCE 10 AA; 1228 MW; 686C88E33054433B CRC64;

Query Match 33.3%; Score 2; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 HS 6
Db      8 HS 9

RESULT 13
Q47561
ID Q47561 PRELIMINARY; PRT; 10 AA.
AC Q47561;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE Hypothetical 1.1 kDa protein (Fragment).
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=K-12;
RX MEDLINE=94162733; PubMed=7764507;
RA Yamada M., Yanai S., Talkner A.;
RT "Analysis of products of the Escherichia coli genomic genes and
RT regulation of their expressions: an applicable procedure for genomic
RT analysis of other microorganisms.";
RL Biosci. Biotechnol. Biochem. 58:117-120(1994).
DR EMBL; D21143; BAA04679.1; -.
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 10 AA; 1109 MW; 2D1B58B1E87DD733 CRC64;

Query Match 33.3%; Score 2; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GH 2
Db      ||

RESULT 14
Q9UVM2
ID Q9UVM2 PRELIMINARY; PRT; 10 AA.
AC Q9UVM2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Bap2(1).
GN BAP2(1).
OS Schizophyllum commune (Bracket fungus).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Agaricales; Schizophyllaceae; Schizophyllum.
OX NCBI_TaxID=5334;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=12-43;
RA Siebert-Bartholmei C., Kothe E.;
RT "Basidiomycete pheromone recognition studied in yeast.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF102785; AAF03757.1; -.
SQ SEQUENCE 10 AA; 1076 MW; D7D803AEB1E861B2 CRC64;

Query Match 33.3%; Score 2; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GH 2
Db      5 GH 6

RESULT 15
Q9HDS2
ID Q9HDS2 PRELIMINARY; PRT; 10 AA.
AC Q9HDS2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE TrpC polypeptide (Fragment).
GN TRPC.
OS Aspergillus flavus.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5059;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=TX19_21S;
RA Geiser D.M., Dörner J.W., Horn B.W., Taylor J.W.;
RT "The phylogenetics of mycotoxin and sclerotium production in
RT Aspergillus flavus and Aspergillus oryzae.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF261865; AAG16139.1; -.
KW Polypeptide.
FT NON_TER
SQ SEQUENCE 10 AA; 997 MW; 8FEBE4EAB2C72AB5 CRC64;

Query Match 33.3%; Score 2; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 HS 6
Db      8 HS 9

RESULT 16
P82223
ID P82223 PRELIMINARY; PRT; 10 AA.
AC P82223;

```

01-OCT-2001 (TRENBLrel. 18, Created)
 01-OCT-2001 (TRENBLrel. 18, Last sequence update)
 01-OCT-2001 (TRENBLrel. 18, Last annotation update)
 Unknown protein from 2D-page (Fragment).
 Bombyx mori (Silk moth).
 Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
 Bombycidae; Bombyx.
 NCBI_TaxID=7091;
 [1]
 SEQUENCE.
 STRAIN=XINHANG X KEMING; TISSUE=Body wall, and Fat body;
 MEDLINE=21177481; PubMed=11280994;
 Zhong B.X.;
 "Protein database for several tissues derived from five instar of
 silkworm.";
 I Chuan Heueh Pao 28:217-224(2001).
 NON_TER 10 10
 SEQUENCE 10 AA; 1054 MW; D0F722C325B1F1B2 CRC64;
 Query Match 33.3%; Score 2; DB 5; Length 10;
 Best Local Similarity 100.0%; Pred.No. 1.1e+04;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Y 5 HS 6
 ||
 b 5 HS 6

RESULT 17
 ID P82224 PRELIMINARY; PRT; 10 AA.
 AC P82224;
 DT 01-OCT-2001 (TRENBLrel. 18, Created)
 DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)
 DT 01-OCT-2001 (TRENBLrel. 18, Last annotation update)
 DE Unknown protein from 2D-page (Fragment).
 DE Bombyx mori (Silk moth).
 DE Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
 DE Neoptera; Endopterygota; Lepidoptera; Glossata; Bombycoidea;
 DE Bombycidae; Bombyx.
 DE NCBI_TaxID=7091;
 DE [1]
 DE SEQUENCE.
 DE STRAIN=XINHANG X KEMING; TISSUE=Body wall, and Fat body;
 DE MEDLINE=21177481; PubMed=11280994;
 DE Zhong B.X.;
 DE "Protein database for several tissues derived from five instar of
 DE silkworm.";
 DE I Chuan Heueh Pao 28:217-224(2001).
 DE NON_TER 10 10
 DE SEQUENCE 10 AA; 1064 MW; D77CBF25B1F1B2CD CRC64;
 Query Match 33.3%; Score 2; DB 5; Length 10;
 Best Local Similarity 100.0%; Pred.No. 1.1e+04;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Y 5 HS 6
 ||
 b 6 HS 7

RESULT 18
 ID P81899 PRELIMINARY; PRT; 10 AA.
 AC P81899;
 DT 01-MAR-2001 (TRENBLrel. 16, Created)
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
 DE Peptide-N4-(N-acetyl-beta-glucosaminyl)asparagine amidase A, large
 DE chain (Subunit A) (EC 3.5.1.52) (PNGase A) (Glycopeptide N-
 DE glycosidase) (N-glycanase) (Fragment).
 DE Prunus dulcis (Almond) (Prunus amygdalus).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
 OC NCBI_TaxID=3755;
 RN [1]
 RP SEQUENCE, AND CHARACTERIZATION.
 RX PubMed=9523720;
 RA Altmann F., Paschinger K., Dalik T., Vorauer K.;
 RT "Characterisation of peptide-N4-(N-acetyl-beta-glucosaminyl)asparagine
 RT amidase A and its N-glycans.";
 RL Eur. J. Biochem. 252:118-123(1998).
 CC -!- CATALYTIC ACTIVITY: HYDROLYSIS OF AN N4-(ACETYL-BETA-D-
 CC GLUCOSAMINYL)ASPARAGINE RESIDUE IN WHICH THE N-ACETYL-D-
 CC (SUBSTITUTED) N-ACETYL-BETA-D-GLUCOSAMINYLAMINE AND THE PEPTIDE
 CC CONTAINING AN ASPARTIC RESIDUE.
 CC -!- SUBUNIT: HETERODIMER OF A LARGE AND A SMALL CHAIN.
 CC -!- PTM: IS HIGHLY GLYCOSYLATED AND IS RESISTANT AGAINST SELF-
 CC DEGLYCOSYLATION.
 CC -!- MASS SPECTROMETRY: MW=54182; METHOD=MALDI.
 KW Hydrolyase; Glycoprotein.
 FT NON_TER 10 10
 SQ SEQUENCE 10 AA; 1106 MW; 95F6BF65B1FB5865 CRC64;
 Query Match 33.3%; Score 2; DB 10; Length 10;
 Best Local Similarity 100.0%; Pred.No. 1.1e+04;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 HS 6
 ||
 DB 6 HS 7

RESULT 19
 ID Q85563 PRELIMINARY; PRT; 10 AA.
 AC Q85563;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE Env-mos fusion protein (fragment).
 OS Moloney murine leukemia virus.
 OC Viruses; Retroid viruses; Retroviridae; Gammaretrovirus.
 OC NCBI_TaxID=11801;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=82196891; PubMed=6281735;
 RA Donoghue D.J., Hunter T.;
 RT "A generalized method of subcloning DNA fragments by restriction site
 RT reconstruction: Application to sequencing the amino-terminal region of
 RT the transforming gene of Gazdar murine sarcoma virus.";
 RL Nucleic Acids Res. 10:2549-2564(1982).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83164305; PubMed=6300424;
 RA Donoghue D.J., Hunter T.;
 RT "Recombination junctions of variants of Moloney murine sarcoma virus:
 RT Generation and divergence of a mammalian transforming gene.";
 RL J. Virol. 45:607-617(1983).
 DR EMBL; K03105; AAA46491.1; -.
 FT NON_TER 10 10
 SQ SEQUENCE 10 AA; 1062 MW; F9ECFCBEA771B591 CRC64;
 Query Match 33.3%; Score 2; DB 15; Length 10;
 Best Local Similarity 100.0%; Pred.No. 1.1e+04;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 HS 6
 ||
 DB 3 HS 4

RESULT 20

```

Q99007
ID Q99007 PRELIMINARY; PRT; 5 AA.
AC Q99007;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Alpha amylase (Fragment).
GN AMY1 GENE.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticaceae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91329704; PubMed=1831055;
RA Jacobsen J.V., Close T.J.;
RT "Control of transient expression of chimaeric genes by gibberellic
RT acid and abscisic acid in protoplasts prepared from mature barley
RT aleurone layers.";
RL Plant Mol. Biol. 16:713-721(1991).
DR EMBL; X54643; CAA38455.1; -.
FT NON_TER 5
SQ SEQUENCE 5 AA; 600 MW; 61E3344D6F00000 CRC64;

Query Match 16.7%; Score 1; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 H 2
|
DB 5 H 5

RESULT 21
ID P82181 PRELIMINARY; PRT; 6 AA.
AC P82181;
DT 01-JUN-2000 (TrEMBLrel. 14, Created)
DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Chloroplast 50S ribosomal protein L10 beta (Fragment).
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE.
RX STRAIN=cv. ALVARO; TISSUE=Leaf;
RX MEDLINE=20435798; PubMed=10874046;
RA Yamaguchi K., Subramanian A.R.;
RT "The plastid ribosomal proteins. Identification of all the proteins in
RT the 50 S subunit of an organelle ribosome (chloroplast).";
RL J. Biol. Chem. 275:28466-28482(2000).
CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC -!- MASS SPECTROMETRY: MW=10477.0; METHOD=ELECTROSPRAY.
CC -!- MASS SPECTROMETRY: MW=10495; METHOD=MALDI.
CC -!- MISCELLANEOUS: S19 ALPHA AND BETA FORMS DIFFER IN PI. S19 BETA
CC FORM IS THE MINOR BASIC FORM.
CC -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 12 KDA.
CC -!- SIMILARITY: BELONGS TO THE S19P FAMILY OF RIBOSOMAL PROTEINS.
DR InterPro; IPR002222; Ribosomal_S19.
DR Pfam; PF00203; Ribosomal_S19; PARTIAL.
DR PRINTS; PR00975; RIBOSOMALS19; PARTIAL.
DR PROSITE; PS00323; RIBOSOMAL_S19; PARTIAL.
KW Ribosomal protein; Chloroplast; rRNA-binding.
FT NON_TER 6
SQ SEQUENCE 6 AA; 732 MW; 63333735A411C000 CRC64;

Query Match 16.7%; Score 1; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
|
DB 3 S 3

RESULT 23
ID P82182 PRELIMINARY; PRT; 6 AA.
AC P82182;
DT 01-JUN-2000 (TrEMBLrel. 14, Created)
DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Chloroplast 50S ribosomal protein L10 gamma (Fragment).
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE.
RX STRAIN=cv. ALVARO; TISSUE=Leaf;
RX MEDLINE=20435798; PubMed=10874046;
RA Yamaguchi K., Subramanian A.R.;
RT "The plastid ribosomal proteins. Identification of all the proteins in
RT the 50 S subunit of an organelle ribosome (chloroplast).";
RL J. Biol. Chem. 275:28466-28482(2000).

```

CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
 CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
 CC -1- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.
 CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
 DR InterPro: IPR001790; Ribosomal L10.
 DR InterPro: IPR002363; Ribosomal_L10eub.
 DR Pfam: PF04666; Ribosomal_L10; PARTIAL.
 DR PROSITE: PS01109; RIBOSOMAL L10; PARTIAL.
 DR RIBOSOMAL protein; Chloroplast; rRNA-binding.
 KW NON TER 6
 FT SEQUENCE 6 AA; 675 MW; 6321B415B05DB000 CRC64;

Query Match 16.7%; Score 1; DB 10; Length 6;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DY 6 S 6
 |
 DB 3 S 3

RESULT 24

ID Q8KMS3 PRELIMINARY; PRT; 7 AA.
 AC Q8KMS3
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Putative MerR2 protein.
 GN MERR2.
 OS Klebsiella sp. LS13-39.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Klebsiella.
 CX NCBI_TaxID=143776;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LS13-39;
 RX MEDLINE=21604134; PubMed=11763242;
 RA Mindlin S.Z., Kholodii G.Y., Gorlenko Z.M., Minakhina S.V.,
 RA Minakhin L.S., Kaliyeva E.S., Kopteva A.V., Petrova M.A.,
 RA Yurieva O.V., Nikiforov V.G.;
 RT "Mercury resistance transposons of Gram-negative environmental
 bacteria and their classification."
 RL Res. Microbiol. 152:811-822(2001).
 DR EMBL: AJ302776; CAC82975.1; -.
 DR EMBL: AJ302776; CAC82975.1; -.
 SQ SEQUENCE 7 AA; 608 MW; 6DC1B5BDD87DD6F0 CRC64;

Query Match 16.7%; Score 1; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DY 1 G 1
 |
 DB 3 G 3

RESULT 25

ID Q47505 PRELIMINARY; PRT; 7 AA.
 AC Q47505
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE MCCA protein.
 GN MCCA.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 CX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=9609297; PubMed=8522520;
 RA Gonzalez-Pastor J.E., San Millan J.L., Castilla M.A., Moreno F.;
 RT "Structure and organization of plasmid genes required to produce the
 RT translation inhibitor microcin C7."
 RL J. Bacteriol. 177:7131-7140(1995).
 DR EMBL: X57583; CAA40808.1; -.
 KW Plasmid.
 SQ SEQUENCE 7 AA; 763 MW; 644DD44861B406F0 CRC64;

Query Match 16.7%; Score 1; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DY 1 G 1
 |
 DB 4 G 4

RESULT 26

ID P70804 PRELIMINARY; PRT; 7 AA.
 AC P70804
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Algt protein (Fragment).
 GN ALGT.
 OS Azotobacter vinelandii.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Azotobacter.
 CX NCBI_TaxID=354;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=E; PubMed=9830682;
 RX MEDLINE=96427318; PubMed=9830682;
 RA Renm B.H.A., Ertesvag H., Valla S.;
 RT "A new Azotobacter vinelandii mannuronan C-5-epimerase gene (algG) is
 RT part of an alg gene cluster physically organized in a manner similar
 RT to that in pseudomonas aeruginosa."
 RL J. Bacteriol. 178:5884-5889(1996).
 DR EMBL: X87973; CAA61230.1; -.
 DR NON TER 1
 SQ SEQUENCE 7 AA; 684 MW; 71B5A5A5A2D1AED0 CRC64;

Query Match 16.7%; Score 1; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DY 6 S 6
 |
 DB 4 S 4

RESULT 27

ID Q47029 PRELIMINARY; PRT; 7 AA.
 AC Q47029
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
 DE Aad A1 protein (Fragment).
 GN Aad A1.
 OS Enterobacter cloacae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Enterobacter.
 CX NCBI_TaxID=550;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94079349; PubMed=8257126;
 RA Rather P.N., Mann P.A., Mierzwa R., Hare R.S., Miller G.H., Shaw K.J.;
 RT "Analysis of the aac(3)-Via gene encoding a novel 3-N-
 RT acetyltransferase."
 RL Antimicrob. Agents Chemother. 37:2074-2079(-993).

DR EMBL; M88012; AAA16193.1; --
 FT NON TER 1
 SQ SEQUENCE 7 AA; 744 MW; 633862D2C321A030 CRC64;

Query Match 16.7%; Score 1; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 DB 6 G 6

RESULT 28
 P72081 PRELIMINARY; PRT; 7 AA.
 ID P72081
 AC P72081
 DT 01-FEB-1997 (TReMBLrel. 02, Created)
 DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE 3'-methylcephem hydroxylase (Fragment).
 GN CEF3.
 OS Nocardia lactamdurans.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Pseudonocardineae; Pseudonocardaceae; Amycolatopsis.
 OX NCBI_TaxID=1913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96009872; PubMed=7557411;
 RA Coque J., Perez-Blarena F.J., Enguita F.J., Fuente J.L., Martin J.F.,
 RA Liras P.;
 RT "Characterization of the cmcH genes of Nocardia lactamdurans and
 RT Streptomyces clavuligerus encoding a functional 3'-hydroxymethylcephem
 RT O-carbamoyltransferase for cephamycin biosynthesis.";
 RL Gene 162:21-27(1995).
 DR EMBL; Z21682; CAA79797.1; --
 FT NON TER 1
 SQ SEQUENCE 7 AA; 746 MW; 75A1B2CDD1E681C0 CRC64;

Query Match 16.7%; Score 1; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 H 2
 DB 3 H 3

RESULT 29
 O07354 PRELIMINARY; PRT; 7 AA.
 ID O07354
 AC O07354
 DT 01-JUL-1997 (TReMBLrel. 04, Created)
 DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE NiK (Fragment).
 GN NIFK.
 OS Synechococcus sp. (strain PCC 8801 / RF-1) (Cyanothecae PCC 8801).
 OC Bacteria; Cyanobacteria; Chroococcales; Cyanothecae.
 OX NCBI_TaxID=41431;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RF-1;
 RX MEDLINE=99231861; PubMed=10217509;
 RA Huang T.C., Lin R.F., Chu M.K., Chen H.M.;
 RT "Organization and expression of nitrogen-fixation genes in the aerobic
 RT nitrogen-fixing unicellular cyanobacterium Synechococcus sp. strain
 RT RF-1.";
 RL Microbiology 145:743-753(1999).
 DR EMBL; AF003700; AAC35193.1; --
 FT NON TER 1
 SQ SEQUENCE 7 AA; 849 MW; 7412C72AA9D5B030 CRC64;

Query Match 16.7%; Score 1; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
 DB 2 S 2

RESULT 30
 Q28742 PRELIMINARY; PRT; 7 AA.
 ID Q28742
 AC Q28742
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE Alpha-myosin heavy chain (Fragment).
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84221901; PubMed=6328491;
 RA Friedman D.J., Umeda P.K., Sinha A.M., Hsu H.-J., Jokovic S.,
 RA Rabinowitz M.;
 RT "Characterization of genomic clones specifying rabbit alpha- and beta-
 RT ventricular myosin heavy chains.";
 RL Proc. Natl. Acad. Sci. U.S.A. 81:3044-3048(1984).
 DR EMBL; K01698; AAA31415.1; --
 FT NON TER 1
 SQ SEQUENCE 7 AA; 916 MW; 6B1B1AA1E69326B0 CRC64;

Query Match 16.7%; Score 1; DB 6; Length 7;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 H 2
 DB 4 H 4

RESULT 31
 O99182 PRELIMINARY; PRT; 7 AA.
 ID O99182
 AC O99182
 DT 01-MAY-1999 (TReMBLrel. 10, Created)
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE Cytochrome oxidase I (Fragment).
 GN COI.
 OS Gnatholebias zonatus.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
 OC Cyprinodontiformes; Aplocheilidae; Rivulinae; Gnatholebias.
 OX NCBI_TaxID=135316;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20072928; PubMed=10603257;
 RA Murphy W.J., Thomson J.R., Collier G.E.;
 RT "Phylogeny of the Neotropical killifish family Rivulidae
 RT (Cyprinodontiformes, Aplocheiloidei) inferred from mitochondrial DNA
 RT sequences";
 RL Mol. Phylogenet. Evol. 13:289-301(1999).
 DR EMBL; AF002591; AAD01074.1; --
 KW Mitochondrion.
 FT NON TER 1
 SQ SEQUENCE 7 AA; 899 MW; 672721F6CB572030 CRC64;

Query Match 16.7%; Score 1; DB 8; Length 7;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 H 2
|
Db 5 H 5

RESULT 32

Q8MFY6 Q8MFY6 PRELIMINARY; PRT; 7 AA.
ID Q8MFY6
AC Q8MFY6
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE P8BA (Fragment).
DE P8BA (Fragment).
GN PSBA.
OS Taraxacum (sect. Dioszegia) sp. 4310Hnew.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; campanulids; Asterales; Asteraceae; Cichorioideae;
OC Cichorieae; Taraxacum.
OC NCBI_TaxID=154248;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97351561; PubMed=9207843;
RA Ostler J.H., Olson D.C., Shiu O.Y., Yang S.F.;
RT "Differential induction of seven l-aminocyclopropane-l-carboxylate
synthase genes by elicitor in suspension cultures of tomato
(Lycopersicon esculentum).";
RL Plant Mol. Biol. 34:275-286(1997).
DR EMBL; AY015477; AAK21591.1; -.
KW Chloroplast.
FT NON TER 1
SQ SEQUENCE 7 AA; 675 MW; 687451B5A76DDB70 CRC64;

Query Match 16.7%; Score 1; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
|
Db 4 S 4

RESULT 33

Q95945 Q95945 PRELIMINARY; PRT; 7 AA.
ID Q95945
AC Q95945
DT 01-FEB-1997 (TReMBLrel. 02, Created)
DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 21, Last annotation update)
DE Inside intron 5 (Fragment).
DE Saccharomyces cerevisiae (Baker's yeast).
OC Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OC NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=D273-10B;
MEDLINE=81069885; PubMed=6254986;
RA Bonitz S.G., Coruzzi G., Thalendorf B., Tzagoloff A., Macino G.;
RT "Assembly of the mitochondrial membrane system: Structure and
nucleotide sequence of the gene coding for subunit 1 of yeast
cytochrome oxidase".
RL J. Biol. Chem. 255:11927-11941(1980).
DR EMBL; V00694; CAA24066.1; -.
FT NON TER 1
SQ SEQUENCE 7 AA; 859 MW; 75B7232362CDC460 CRC64;

Query Match 16.7%; Score 1; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
|
Db 7 S 7

RESULT 34

P93233 P93233 PRELIMINARY; PRT; 7 AA.
ID P93233
AC P93233
DT 01-MAY-1997 (TReMBLrel. 03, Created)
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE l-aminocyclopropane-l-carboxylate synthase (EC 4.4.1.14)
DE (Fragment).
GN LS-ACS1B.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamids; Solanales; Solanaceae; Solanum.
OC NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97351561; PubMed=9207843;
RA Ostler J.H., Olson D.C., Shiu O.Y., Yang S.F.;
RT "Differential induction of seven l-aminocyclopropane-l-carboxylate
synthase genes by elicitor in suspension cultures of tomato
(Lycopersicon esculentum).";
RL Plant Mol. Biol. 34:275-286(1997).
DR EMBL; U75692; AAC49682.1; -.
KW Lyase.
FT NON TER 1
SQ SEQUENCE 7 AA; 828 MW; 7B412C7377415D0 CRC64;

Query Match 16.7%; Score 1; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
|
Db 1 S 1

RESULT 35

O49223 O49223 PRELIMINARY; PRT; 7 AA.
ID O49223
AC O49223
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE HMG-1-like protein (Fragment).
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosidis I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OC NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Essex; TISSUE=Root;
MEDLINE=91367679; PubMed=1891369;
RA Laux T., Goldberg R.B.;
RT "A plant DNA binding protein shares highly conserved sequence motifs
with HMG-box proteins".
RL Nucleic Acids Res. 19:4769-4769(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Essex; TISSUE=Root;
RA Mahalingam R., Knap H.T.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF047050; AAC03556.1; -.
FT NON TER 1
SQ SEQUENCE 7 AA; 850 MW; 6AAAAAB378637810 CRC64;

```
Query Match      16.7%; Score 1; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels
```

Qy	1	1	1	1
Db	1	1	1	1

RESULT 36
Q9C5B3
ID Q9C5B3 PRELIMINARY; PRT: 7 AA.

01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
Hypothetical 0.7 kDa protein (Fragment).
DID1_10A-2B.
Arabidopsis thaliana (Mouse-ear cress).
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
eurosids II; Brassicales; Brassicaceae; Arabidopsids.
NCBI_faxID=3702;
[1]
SEQUENCE FROM N.A.
TISU01:Root;
MEDLINE=21171025; PubMed=11277426;
Vercateren I., Van Der Schueren E., Van Montagu M., Gheysen G.;
"Arabidopsis thaliana genes expressed in the early compatible
interaction with root-knot nematodes."; Mol. Plant Microbe Interact. 14:288-299 (2001).
EMBL; AJ286350; CAB71014.2; -.

```
Query Match      16.7%; Score 1; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches      1; Conservative 0; Mismatches 0; Indels
```

Qy	6 5 6
Db	1 5 1

RESULT 37
Q8K3H6
ID Q8K3H6 PRELIMINARY; PRT; 7 AA.

AC	GenShare;
AD	01-OCT-2002 (TrEMBLrel. 22, Created)
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DD	01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE	Collagenase-3 (Fragment).
DN	MMP13.
GN	Rattus norvegicus (Rat).
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX	NCBI_TaxID=10116;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=Sprague-Dawley;
RT	Prietto S.M., Lyons J.G.;
RA	"Intron 1 of Rattus norvegicus MMP13.";
RRL	Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR	EMBL; AY101357; AAU551172.1.; -
FD	NON_TER 1 1
FT	NON_TER 1 1
FT	NON_TER 7 7
SQ	SEQUENCE 7 AA: 907 MW: 63373B51EBDD9A0 CRC64:

```
Query Match      16.7%; Score 1; DB 11; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels
```

$$\begin{array}{c} \text{Qy} \\ 2 \text{ H } 2 \\ | \\ 4 \text{ H } 4 \\ \text{Db} \end{array}$$

RESULT 38
Q63480
ID Q63480 PRELIMINARY; PRT; 7 AA.

Q03460.	01-NOV-1996	(TrEMBLrel. 01, Created)	DT
DT	01-NOV-1996	(TrEMBLrel. 01, Last sequence update)	DT
DT	01-DEC-2001	(TrEMBLrel. 19, Last annotation update)	DT
DE	TR4-NS orphan receptor (Fragment).		DE
TR4.			TR4.
OS	Rattus norvegicus (Rat).		OS
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		OC
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		OC
ON	NCBI_TaxID=10116;		ON
OX	[1]		OX
RP	SEQUENCE FROM N.A.		RP
RP	MEDLINE=96198747; PubMed=8612486;		RP
RA	Yoshikawa T.; Makino S.; Gao X.M.; Xing G.Q.; Chuang D.M.;		RA
RA	Detera-Wadleigh S.D.;		RA
RT	"Splice variants of rat TR4 orphan receptor: differential expression		RT
RT	of novel sequences in the 5'-untranslated region and C-terminal		RT
RT	domain.";		RT
RL	Endocrinology 137:1562-1571(1996).		RL
DR	EMBL; U59125; AAB02827.1; -.		DR
KW	Receptor.		KW
FT	NON_TER		FT
FT	1		FT
SQ	SEQUENCE 7 AA; 758 MW; 672AAB7864005350 CRC64;		SQ

Query Match 16.7%; Score 1; DB 11; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps

Qy 1 G 1
Db 4 G 4

RESULT 39
Q55184
ID Q55184
PRELIMINARY:
PRT: 7 AA.

AC	055184;
AD	01-JUN-1998 (TrEMBLrel. 06, Created)
AE	01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
AF	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
AG	Orphan receptor TR4-NS (fragment).
AH	TR4.
AI	Rattus norvegicus (Rat).
AJ	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AK	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus;
AL	NCBI_TaxID=10116;
AM	[1]_TaxID=10116;
AN	SEQUENCE FROM N.A.
AO	STRAIN=Sprague-Dawley;
AP	MEDLINE=96198747; PubMed=8612486;
AQ	Yoshikawa T., Makino S., Gao X.M., Xing G.Q., Chuang D.M.,
AR	Detera-Wadleigh S.D.;
AS	"Splice variants of rat TR4 orphan receptor; differential expression
AT	of novel sequences in the 5'-untranslated region and C-terminal
AV	domain.;"
AW	Endocrinology 137:1562-1571(1996).
AX	[2]
AY	SEQUENCE FROM N.A.
AZ	STRAIN=Sprague-Dawley;
BA	MEDLINE=96299786; PubMed=8661150;
BB	Yoshikawa T., Dupont B.R., Leach R.J., Detera-Wadleigh S.D.;
BC	"New variants of the human and rat nuclear hormone receptor, TR4:
BD	expression and chromosomal localization of the human gene.;"
BE	Genomics 35:361-366(1996).

```
DR EMBL; U59454; AAB91433.1; -.
KW Receptor.
FT NON_TER 1 1
SQ SEQUENCE 7 AA; 663 MW; 6DDAA8787EB05350 CRC64;

Query Match
 16.7%; Score 1; DB 11; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
Db 4 G 4

RESULT 40
ID Q63668 PRELIMINARY; PRT; 7 AA.
AC Q63668;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ORF2 protein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Kidney;
RC MEDLINE=95396550; PubMed=7667072;
RA Mandon B., Bellanger A.C., Elalouf J.M.;
RA "Inverse-PCR-mediated cloning of the promoter for the rat vasopressin
RT V2 receptor gene";
RT Pfluegers Arch. 430:12-18(1995).
RRL EMBL; X83264; CAA58237.1; -.
SQ SEQUENCE 7 AA; 703 MW; 75A767287DC6D6F0 CRC64;

Query Match
 16.7%; Score 1; DB 11; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
Db 4 G 4

RESULT 41
ID Q66205 PRELIMINARY; PRT; 7 AA.
AC Q66205;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Membrane protein (1 is 3rd base in codon) (Fragment).
OS Transmissible gastroenteritis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OC NCBI_TaxID=11149;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FS772/70;
RC MEDLINE=88216185; PubMed=2835592;
RA Britton P., Carmenes R.S., Page K.W., Garwes D.J., Parra F.;
RA "Sequence of the Nucleoprotein Gene from a Virulent British Field
RT Isolate of Transmissible Gastroenteritis Virus and its Expression in
RT Saccharomyces Cerevisiae.";
RRL Mol. Microbiol. 2:89-99(1988).
RRL EMBL; Y00542; CAA68606.1; -.
FT NON_TER 1 1
SQ SEQUENCE 7 AA; 869 MW; 72C691E727233B70 CRC64;

Query Match
 16.7%; Score 1; DB 12; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;

QY 1 G 1
Db 4 G 4

us-10-057-890a-15.oligo.rsp

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 H 2
Db 5 H 5

RESULT 42
ID Q8JJ20 PRELIMINARY; PRT; 7 AA.
AC Q8JJ20;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Extracellular fatty acid binding protein (Fragment).
GN EXPABP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA Wang Q., Li N., Li H.;
RT "Cloning and sequencing of 3' UTR of EXPABP gene in chicken.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF487519; AAL96665.1; -.
FT NON_TER 1 1
SQ SEQUENCE 7 AA; 780 MW; 72CB1AB2D5BEBB70 CRC64;

Query Match
 16.7%; Score 1; DB 13; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
Db 3 S 3

RESULT 43
ID O42564 PRELIMINARY; PRT; 7 AA.
AC O42564;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE Truncated voltage-gated sodium channel alpha subunit (Fragment).
GN SCN8A.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontoidea; Tetraodontidae; Takifugu.
OC NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=97442476; PubMed=9295353;
RA Plummer N.W., McBurney M.W., Meisler M.H.;
RT "Alternative splicing of the sodium channel SCN8A predicts a truncated
RT two-domain protein in fetal brain and non-neuronal cells.";
RRL J. Biol. Chem. 272:24008-24015(1997).
DR EMBL; U97673; AAB80916.1; -.
KW Ionic channel.
FT NON_TER 1 1
SQ SEQUENCE 7 AA; 730 MW; 75B72EA2C73772A0 CRC64;

Query Match
 16.7%; Score 1; DB 13; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
Db 3 S 3
```

```

Db          7 S 7

RESULT 44
Q07624      PRELIMINARY;      PRT;      7 AA.
AC Q07624;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE UORF1.
OS Rous sarcoma virus (strain Prague C).
OC Viruses; Retrov. retroviridae; Alpharetrovirus.
OX NCBI_TaxID=11889;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93010967; PubMed=1327749;
RA Donze O., Spahr P.F.;
RT "Role of the open reading frames of Rous sarcoma virus leader RNA in
RT translation and genome packaging.";
RL EMBO J. 11:3747-3757(1992).
DR EMBL: X67587; CAA47862.1; -.
SQ SEQUENCE 7 AA; 672 MW; 776045A7687DD6F0 CRC64;

Query Match      16.7%; Score 1; DB 15; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 G 1
Db 3 G 3

RESULT 45
Q45615      PRELIMINARY;      PRT;      8 AA.
AC Q45615;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE GutB protein (Fragment).
GN GUTB
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94253000; PubMed=8195086;
RA Ye R., Wong S.L.;
RT "Transcriptional regulation of the Bacillus subtilis glucitol
RT dehydrogenase gene.";
RL J. Bacteriol. 176:3314-3320(1994).
DR EMBL: L16626; AAA20875.1; -.
FT NON TER 8
SQ SEQUENCE 8 AA; 927 MW; FD56C772D1A1F1A6 CRC64;

Query Match      16.7%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 H 2
Db 3 H 3

RESULT 46
Q9AGP4      PRELIMINARY;      PRT;      8 AA.
AC Q9AGP4;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Shigella dysenteriae.
OS Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

```

```

DE Serine hydroxymethyltransferase (Fragment).
GN GLYA.
OS Arthrobacter sp. 11N.
OC Bacteria; Actinobacteridae; Actinomycetales;
OC Micrococccineae; Micrococcaceae; Arthrobacter.
OX NCBI_TaxID=153502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=11N;
RA Meskys R., Harris R.J., Casalte V., Baaran J., Scrutton N.S.;
RT "Genetic organization of the genes involved in dimethylglycine and
RT sarcosine degradation in Arthrobacter spp.: implications for glycine
RT betaine catabolism.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF329478; AAK16486.1; -.
KW Methyltransferase; Transferase.
FT NON TER 1
SQ SEQUENCE 8 AA; 898 MW; 6B1870533372457 CRC64;

Query Match      16.7%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 G 1
Db 7 G 7

RESULT 47
Q9R3X0      PRELIMINARY;      PRT;      8 AA.
AC Q9R3X0;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit
DE (Fragment).
GN RBCL.
OS Planktothrix rubescens.
OC Bacteria; Cyanobacteria; Oscillatoriales; Planktothrix.
OX NCBI_TaxID=59512;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BC-PLA 9316, and BC-PLA 9303;
RX MEDLINE=20005589; PubMed=10537197;
RA Beard S.J., Handley B.A., Hayes P.K., Walsby A.E.;
RT "The diversity of gas vesicle genes in Planktothrix rubescens from
RT Lake Zurich.";
RL Microbiology 145:2757-2768(1999).
DR EMBL: AJ132249; CAB59537.1; -.
DR EMBL: AJ132248; CAB59534.1; -.
FT NON TER 1
SQ SEQUENCE 8 AA; 957 MW; 33D1AAA685BB19CB CRC64;

Query Match      16.7%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6
Db 4 S 4

RESULT 48
Q9R5R2      PRELIMINARY;      PRT;      8 AA.
AC Q9R5R2;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE 31,000 DA product of ORF8 (Fragment).
OS Shigella dysenteriae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

```

OC Enterobacteriaceae; Shigella.
 OX NCBI_TaxID=622;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=92085268; PubMed=1660923;
 RA Polard P., Prere M.F., Chandler M., Payet O.;
 "Programmed translational frameshifting and initiation at an AUC codon
 RT in gene expression of bacterial insertion sequence IS911.";
 RL J. Mol. Biol. 222:465-477(1991).
 RFT NON_TER 8
 SQ SEQUENCE 8 AA; 902 MW; FE2DCAPF586AE336 CRC64;

Query Match 16.7%; Score 1; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 G 1
 Db 4 G 4

RESULT 49

ID O68485 PRELIMINARY; PRT; 8 AA.
 AC O68485;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Aminoglycoside 3'-O-nucleotidyltransferase (Fragment).
 GN AADAL.
 OS Klebsiella pneumoniae.
 OC Plasmid pLQ1000.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Klebsiella.
 OX NCBI_TaxID=573;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K1;
 RX MEDLINE=98287600; PubMed=9624504;
 RA Centron D., Roy P.H.;
 RT "Characterization of the 6'-N-aminoglycoside acetyltransferase gene
 RT aac(6')-Iq from the integron of a natural multiresistance plasmid.";
 RL Antimicrob. Agents Chemother. 42:1506-1508(1998).
 DR EMBL: AF047556; AAC25501.1; -.
 KW Plasmid; Transferase.
 FT NON_TER 8
 SQ SEQUENCE 8 AA; 878 MW; EFC1B5A2CB1DD056 CRC64;

Query Match 16.7%; Score 1; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6
 Db 6 S 6

RESULT 50

ID Q93SR0 PRELIMINARY; PRT; 8 AA.
 AC Q93SR0;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Beta-lactamase repressor BlaI (Fragment).
 GN BLAI.
 OS Staphylococcus epidermidis.
 OC Plasmid pSt6.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=1282;
 RN [1]
 RP SEQUENCE FROM N.A.
 RFT STRAIN=6;

RA Sidhu M.S., Heir E., Sorum H., Holck A.L.;
 RT "Genetic linkage between quaternary ammonium compound and beta-lactam
 RT resistance in *Staphylococcus* isolated from food.";
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY028779; AAK38453.1; -.
 KW Plasmid.
 FT NON_TER 1
 SQ SEQUENCE 8 AA; 930 MW; 4E3325B05AA44720 CRC64;

Query Match 16.7%; Score 1; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6
 Db 6 S 6

RESULT 51

ID Q8RJ10 PRELIMINARY; PRT; 8 AA.
 AC Q8RJ10;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE DNA replication initiation protein (Fragment).
 GN DNAA-LIKE.
 OS Streptomyces caespitosus.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=53502;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC27422;
 RA Ma W., Mao X., Jie L., Qin Z.J., Jiang W.H., Chiao R.S., Zhao G.P.;
 RT "Cloning, sequence analysis and function analysis of the replication
 RT origin oric of *Streptovirgillum caespitosus* ATCC27422.";
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ458440; CAD30324.1; -.
 KW NON_TER 1
 SQ SEQUENCE 8 AA; 915 MW; 047443304040451A7 CRC64;

Query Match 16.7%; Score 1; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 G 1
 Db 8 G 8

RESULT 52

ID Q44463 PRELIMINARY; PRT; 8 AA.
 AC Q44463;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE *Agrobacterium tumefaciens* Ti plasmid virD3 and virD4 genes
 DE (Fragment).
 OS Rhizobiales (rhizobacteria).
 OG Plasmid Ti.
 OC Bacteria; Proteobacteria; Alphaproteobacteria.
 OX NCBI_TaxID=356;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88015611; PubMed=3658701;
 RA Portner S.G., Yanofsky M.F., Nester E.W.;
 RT "Molecular characterization of the virD operon from *Agrobacterium*
 RT *tumefaciens*.";
 RL Nucleic Acids Res. 15:7503-7517(1987).
 DR EMBL: X06045; CAA29439.1; -.
 KW Plasmid.

FT NON TER 8 8
SQ SEQUENCE 8 AA; 887 MW; F9F2C325B33861A6 CRC64;

Query Match 16.7%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
DB 3 G 3

RESULT 53
P77556 PRELIMINARY; PRT; 8 AA.

AC P77556;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)
DE TRAY (Fragment).
GN TRAY.
OS Escherichia coli.
OG Plasmid IncFII R1.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ECOR11;
RX MEDLINE=9640G908; PubMed=8807284;
RA Boyd E.F., Hill C.W., Rich S.M., Hartl D.L.;
RT "Mosaic structure of plasmids from natural populations of Escherichia coli";
RL Genetics 143:1091-1100(1996).
DR EMBL; U50661; AAC44245.1; -
DR EMBL; U50650; AAC44234.1; -
DR EMBL; U50651; AAC44235.1; -
DR EMBL; U50652; AAC44236.1; -
DR EMBL; U50653; AAC44237.1; -
DR EMBL; U50654; AAC44238.1; -
DR EMBL; U50655; AAC44239.1; -
DR EMBL; U50656; AAC44240.1; -
DR EMBL; U50657; AAC44241.1; -
DR EMBL; U50658; AAC44242.1; -
DR EMBL; U50659; AAC44243.1; -
DR EMBL; U50660; AAC44244.1; -
KW Plasmid.
FT NON TER
SQ SEQUENCE 8 AA; 834 MW; D335A5B0544735A1 CRC64;

Query Match 16.7%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
DB 2 S 2

RESULT 54
Q9ZEZ9 PRELIMINARY; PRT; 8 AA.

AC Q9ZEZ9;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE 2-isopropylmalate synthase (EC 4.1.3.12) (Fragment).
GN LEUA.
OS Buchnera aphidicola.
OG Plasmid pBRcl.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=9;

[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=99028904; PubMed=9812361;
RX Silva F.J., Van Ham R.C.H.J., Sabater B., Latorre A.;
RT "Structure and evolution of the leucine plasmids carried by the endosymbiont (Buchnera aphidicola) from aphids of the family Aphididae";
RL FEMS Microbiol. Lett. 168:43-49(1998).
DR EMBL; AJ068874; CAA07290.1; -
KW Lyase; Plasmid.
FT NON TER
SQ SEQUENCE 8 AA; 917 MW; FE205042D6C5B446 CRC64;

Query Match 16.7%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
DB 3 S 3

RESULT 55
O092S8 PRELIMINARY; PRT; 8 AA.

AC O092S8;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NifH (Fragment).
GN NifH.
OS Synchococcus sp. (strain PCC 8801 / RF-1) (Cyanothecae PCC 8801).
OC Bacteria; Cyanobacteria; Chroococcales; Cyanothecae.
OX NCBI_TaxID=41431;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RF-1;
RX MEDLINE=99231861; PubMed=10217509;
RA Huang T.C., Lin R.F., Chu M.K., Chen H.M.;
RT "Organization and expression of nitrogen-fixation genes in the aerobic nitrogen-fixing unicellular cyanobacterium Synchococcus sp. strain RF-1";
RL Microbiology 145:743-753(1999).
DR EMBL; AF001780; AAC33369.1; -
FT NON TER
SQ SEQUENCE 8 AA; 985 MW; F16B59CDD046C406 CRC64;

Query Match 16.7%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
DB 8 G 8

RESULT 56
Q8RSR3 PRELIMINARY; PRT; 8 AA.

AC Q8RSR3;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE PduG protein (Fragment).
GN PduG.
OS Lactobacillus collinoides.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=33960;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LWG 18850;
RA Sauvageot N., Gouffi K., Laplace J.M., Auffray Y.;

"Characterisation of the diol dehydratase pdu operon of *Lactobacillus*

collinoides." Lett. 0.0-0(0).

FEMS Microbiol. Lett. 0.0-0(0).

EMBL; AJ297723; CAD01093.1; -.

NON_TER 8

SEQUENCE 8 AA; 882 MW; ECA40B05BAB5BDD6 CRC64;

Query Match 16.7%; Score 1; DB 2; Length 8;

Best Local Similarity 100.0%; Pred. No. 8.3e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6

Db 3 S 3

RESULT 57

Q9X3K1 PRELIMINARY; PRT; 8 AA.

AC Q9X3K1; (Created)

DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)

DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

DE Cytochrome b (Fragment).

Query Match 16.7%; Score 1; DB 2; Length 8;

Best Local Similarity 100.0%; Pred. No. 8.3e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6

Db 1 S 1

RESULT 59

Q56140 PRELIMINARY; PRT; 8 AA.

AC Q56140; (Created)

DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)

DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)

DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).

DE STP6 protein (Fragment).


```

Db          5 H 5
RESULT 61
Q56429      PRELIMINARY;      PRT;      8 AA.
AC Q56429;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GAPDH (Fragment)
OS Thermus thermophilus.
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=274;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HB-8;
RA MEDLINE=89025722; PubMed=3052437;
RA Bowen D., Littlechild J.A., Fothergill J.E., Watson H.C., Hall L.;
RT "Nucleotide sequence of the phosphoglycerate kinase gene from the
RT extreme thermophile, Thermus thermophilus."
RL Biochem. J. 254:509-517(1988).
DR EMBL: X12464; CAA31005.1; -.
FT NON_TER 1
SQ SEQUENCE 8 AA; 885 MW; 33C87333732C72B CRC64;

Query Match      16.7%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 G 1
Db 7 G 7

RESULT 62
Q9R9E0      PRELIMINARY;      PRT;      8 AA.
AC Q9R9E0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE Stage V sporulation protein E (Fragment).
GN SPOVE.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=93003529; PubMed=1391053;
RA Henriques A.O., de Lencastre H., Pigott P.J.;
RT "A Bacillus subtilis morphogene cluster that includes spove is
RT homologous to the mra region of Escherichia coli."
RL Biochimie 74:735-748(1992).
DR EMBL: X64258; CAA45556.1; -.
FT NON_TER 8
SQ SEQUENCE 8 AA; 893 MW; EE75A1A33321B1A6 CRC64;

Query Match      16.7%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6
Db 7 S 7

RESULT 63
Q93SP2      PRELIMINARY;      PRT;      8 AA.
ID Q93SP2
AC Q93SP2;

```

```

DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Quaternary ammonium compound-resistance protein QacEdelta
DE (Fragment).
GN QACEDELTA1.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YMC704; TRANSPOSON=class I integron;
RA Lee K., Chong Y., Yum J.H., Yong D., Livermore D.M.;
RT "VIM-2 metallo-beta-lactamase gene-containing integron in a
RT Pseudomonas aeruginosa clinical isolate."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY029772; AAK50441.1; -.
FT NON_TER 8
SQ SEQUENCE 8 AA; 920 MW; FE2729C732C87336 CRC64;

Query Match      16.7%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 G 1
Db 3 G 3

RESULT 64
Q45889      PRELIMINARY;      PRT;      8 AA.
AC Q45889;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HA-11 protein (Fragment).
GN HA-11.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=667Ab;
RX MEDLINE=96210012; PubMed=8631890;
RA Hutson R.A., Zhou Y., Collins M.D., Johnson E.A., Hatheway C.L.,
RA Sugiyama H.;
RT "Genetic characterisation of Clostridium botulinum type A containing
RT silent type B neurotoxin gene sequences."
RL J. Biol. Chem. 271:10786-10792(1996).
DR EMBL: X87850; CAA61130.1; -.
FT NON_TER 8
SQ SEQUENCE 8 AA; 982 MW; FE29D1B40B02D5B6 CRC64;

Query Match      16.7%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6
Db 2 S 2

RESULT 65
Q9RSR0      PRELIMINARY;      PRT;      8 AA.
ID Q9RSR0
AC Q9RSR0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 11,500 DA product of ORFA (Fragment).

```

OS Shigella dysenteriae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC Enterobacteriaceae; Shigella.
[1] NCBI_TaxID=622;
RN SEQUENCE.
RX MEDLINE=92085268; PubMed=1660923;
RA Polard P., Prere M.P., Chandler M., Fayet O.;
*Programmed translational frameshifting and initiation at an AUU codon
in gene expression of bacterial insertion sequence IS911.";
RT J. Mol. Biol. 222:465-477(1991).
DE NON TER 8
FT SEQUENCE 8 AA; 935 MW; F4C5B9C45AE33336 CRC64;
SQ

Query Match 16.7%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
DB 7 S 7

RESULT 66
Q9R9C2 PRELIMINARY; PRT; 8 AA.
AC Q9R9C2
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE Plasmid cp32-3, possible partition proteins (Fragment).
DE Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid cp32-3
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RX SEQUENCE FROM N.A.
RA STRAIN=B31;
MEDLINE=98361033; PubMed=9695920;
STEVENS B., CASJENS S., ROSA P.;
*Evidence of past recombination events among the genes encoding the
Erp antigens of Borrelia burgdorferi.";
RT Microbiology 144:1869-1879(1998).
DE EMBL; AF022480; AAC35445.1; -.
OG Plasmid.
OC NON TER.
OX SEQUENCE 8 AA; 985 MW; E8B41B41A735B446 CRC64;
SQ

Query Match 16.7%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
DB 3 S 3

RESULT 67
Q47273 PRELIMINARY; PRT; 8 AA.
AC Q47273
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Orf33, orf151, orf56, orf96, rus, orf45, orf127, and mmpC
genes.
DE Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RX SEQUENCE FROM N.A.
RA STRAIN=K-12;
MEDLINE=96196428; PubMed=8648624;
RA Mahdi A.A., Sharples G.J., Mandal T.N., Lloyd R.G.;
*Holliday Junction Resolvases encoded by homologous rusa genes in
RT Escherichia coli K-12 and Phage 32.";
RL J. Mol. Biol. 257:561-573(1996).
DR EMBL; X92587; CAA63323.1; -.
SQ SEQUENCE 8 AA; 935 MW; FD4041A05B05B336 CRC64;

Query Match 16.7%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
DB 3 S 3

RESULT 68
Q9R4M3 PRELIMINARY; PRT; 8 AA.
AC Q9R4M3
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE COBI=BACTERIAL sex pheromone.
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RX SEQUENCE.
RA MEDLINE=95290767; PubMed=7772836;
RA Nakayama J., Abe Y., Ono Y., Isogai A., Suzuki A.;
*Isolation and structure of the Enterococcus faecalis sex pheromone,
RT COBI, that induces conjugal transfer of the hemolysin/bacteriocin
RT plasmids, pOBI and pYII.";
RL Biosci. Biotechnol. Biochem. 59:703-705(1995).
SQ SEQUENCE 8 AA; 741 MW; 83D87732C732CDC2 CRC64;

Query Match 16.7%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
DB 7 G 7

RESULT 69
Q934S4 PRELIMINARY; PRT; 8 AA.
AC Q934S4
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MerD protein (Fragment).
GN MERD.
OS Thiobacillus ferrooxidans.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Acidithiobacillus.
OX NCBI_TaxID=920;
RN [1]
RX SEQUENCE FROM N.A.
RA STRAIN=G66; TRANSPOSON=Tn5037;
RA Kalyaeva E.S., Kholodii G.Y., Bass I.A., Gorlenko A.M., Yurieva O.V.,
RA Nikiforov V.G.;
RT "Tn5037, a Tn21-like mercury resistance transposon from Thiobacillus
ferrooxidans.";
RL Russ. J. Genet. 37:972-975(2001).
DR EMBL; AJ251743; CAC69252.1; -.
FT NON TER 8
SQ SEQUENCE 8 AA; 937 MW; E015A2D77B5DD446 CRC64;

Query Match 16.7%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6
|
7 S 7

RESULT 70

P72221 ID P72221 PRELIMINARY; PRT; 8 AA.
AC P72221; 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Alginatase (Fragment).
CN ALY.
OS Pseudomonas sp. (strain OS-ALG-9).
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=86038;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OS-ALG-9;
RA Fujiyama K.;
RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=OS-ALG-9;
RX MEDLINE=93329366; PubMed=8336113;
RA Maki H., Mori A., Fujiyama K., Kinoshita S., Yoshida T.;
RT "Cloning, sequence analysis and expression in Escherichia coli of a
gene encoding an alginatase from Pseudomonas sp. OS-ALG-9.";
RL J. Gen. Microbiol. 139:987-993 (1993).
RW EMBL; D38469; BAA21704.1; -.
KW Lyase.
FT NON TER 1 1
FT NON TER 8 8
SQ SEQUENCE 8 AA; 841 MW; 461DDDC5A5B041BB CRC64;

Query Match 16.7%; Score 1; DB 2; Length 8;

Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6
|
4 S 4

RESULT 71

O85406 ID O85406 PRELIMINARY; PRT; 8 AA.
AC O85406;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 1.0 kDa protein (Fragment).
OS Coxiella burnetii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Coxiellaceae; Coxiella.
OX NCBI_TaxID=777;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nine Mile Phase I;
RX MEDLINE=98348442; PubMed=9683477;
RA Willems H., Jaeger C., Baljer G.;
RT "Physical and genetic map of the obligate intracellular bacterium
Coxiella burnetii.";
RL J. Bacteriol. 180:3816-3822 (1998).
DR EMBL; AF064963; AAD0947.1; -.
KW Hypothetical protein.
FT NON TER 1 1
FT NON TER 8 8
SQ SEQUENCE 8 AA; 993 MW; 046B5AA453772727 CRC64;

Query Match 16.7%; Score 1; DB 2; Length 8;

Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 G 1
|
8 G 8

RESULT 72

Q9R5L7 ID Q9R5L7 PRELIMINARY; PRT; 8 AA.
AC Q9R5L7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE 1,4-beta-D-glucan glucanohydrolase (EC 3.2.1.4) (Fragment).
OS Clostridium thermocellum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OX NCBI_TaxID=1515;
RN [1]
RP SEQUENCE.
RX MEDLINE=92231850; PubMed=1567379;
RA Romaniec M.P., Fauth U., Kobayashi T., Huskisson N.S., Barker P.J.,
RA Demain A.L.;
RT "Purification and characterization of a new endoglucanase from
Clostridium thermocellum.";
RL Biochem. J. 283:69-73 (1992).
SQ SEQUENCE 8 AA; 823 MW; C2CIAB1DD9D1B775 CRC64;

Query Match 16.7%; Score 1; DB 2; Length 8;

Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6
|
1 S 1

RESULT 73

P83158 ID P83158 PRELIMINARY; PRT; 8 AA.
AC P83158;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Photosystem I iron-sulfur center (Photosystem I subunit VII) (9 kDa
polypeptide) (PSI-C) (Fragment).
OS Anabaena sp. (strain L31).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.
OX NCBI_TaxID=29412;
RN [1]
RP SEQUENCE.
RA Apte S.K., Uhlemann E., Schmid R., Altendorf K.;
RL Submitted (OCT-2001) to the SWISS-PROT data bank.
CC -!- FUNCTION: APOPROTEIN FOR THE IRON-SULFUR CENTERS FA AND FB OF THE
CC -!- PHOTOSYSTEM I COMPLEX.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF
CC 'BACTERIAL-TYPE' 4FE-4S FERREDOXINS.
DR InterPro; IPR001450; 4Fe-4S FERREDOXIN.
KW PROSITE; PS00198; 4FE4S FERREDOXIN; PARTIAL.
RW Photosynthesis; Photosystem I; Iron-sulfur; 4Fe-4S; Metal-binding.
FT NON TER 8 8
SQ SEQUENCE 8 AA; 962 MW; C59B505322D1A1F5 CRC64;

Query Match 16.7%; Score 1; DB 2; Length 8;

Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6
|
1 S 1

```

RESULT 74
Q05403 PRELIMINARY; PRT; 8 AA.
Q05403;
01-NOV-1996 (TrEMBLrel. 01, Created)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DNA for ORF's from chromosome XV (Fragment).
Saccharomyces cerevisiae (Baker's yeast).
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
NCBI_TaxID=4932;
[1]
SEQUENCE FROM N.A.
STRAIN=FY1679;
MEDLINE=96021609; PubMed=8533473;
Zumstein E., Pearson B.M., Kallogeropoulos A., Schweizer M.;
"A 29.425 kb segment on the left arm of yeast chromosome XV contains
more than twice as many unknown as known open reading frames.";
Yeast 11:975-986(1995).
EMBL; X83121; CAA58183.1; -.
NCBI_TaxID=4932;
[2]
SEQUENCE 8 AA; 879 MW; 7B5322D2C441E058 CRC64;

Query Match 16.7%; Score 1; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
DB 1 G 1

RESULT 75
O13591 PRELIMINARY; PRT; 8 AA.
O13591;
01-JAN-1998 (TrEMBLrel. 05, Created)
01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
ORF YNL337W (Fragment).
Saccharomyces cerevisiae (Baker's yeast).
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
NCBI_TaxID=4932;
[1]
SEQUENCE FROM N.A.
Obermaier B., Piravandi E., Rinke M.;
Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A.
MIPS;
Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
EMBL; Z71612; CAA96271.2; -.
NCBI_TaxID=4932;
[1]
SEQUENCE 8 AA; 1005 MW; 5CA441E449C9C720 CRC64;

Query Match 16.7%; Score 1; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 H 2
DB 6 H 6

RESULT 76
P87225 PRELIMINARY; PRT; 8 AA.
ID P87225
AC P87225;
01-JUL-1997 (TrEMBLrel. 04, Created)

```

```

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE GIN11 protein (Fragment).
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
NCBI_TaxID=4932;
[1]
SEQUENCE FROM N.A.
RA Wedler H., Wedler E., Scharfe M., Wambutt R.;
Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A.
RA MIPS;
Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z73169; CAA97518.2; -.
FT NON_TER 1
SQ SEQUENCE 8 AA; 1019 MW; 4E21A9C449D5B73B CRC64;

Query Match 16.7%; Score 1; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 S 6
DB 3 S 3

RESULT 77
Q9HDS4 PRELIMINARY; PRT; 8 AA.
AC Q9HDS4;
01-MAR-2001 (TrEMBLrel. 16, Created)
01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE TrpC polyprotein (Fragment).
GN TRPC.
OS Aspergillus flavus.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
NCBI_TaxID=5059;
[1]
SEQUENCE FROM N.A.
RC STRAIN=A55;
Gaiser D.M., Dörner J.W., Horn B.W., Taylor J.W.;
"The phylogenetics of mycotoxin and sclerotium production in
Aspergillus flavus and Aspergillus oryzae.";
Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF261861; AAG16135.1; -.
KW Polyprotein.
FT NON_TER 8
SQ SEQUENCE 8 AA; 807 MW; F3B2C72AB5B87DD6 CRC64;

Query Match 16.7%; Score 1; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
DB 3 G 3

RESULT 78
Q9URB9 PRELIMINARY; PRT; 8 AA.
ID Q9URB9
AC Q9URB9;
01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Aminopeptidase B (Fragment).
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

```

```

OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE.
RX MEDLINE=92088139; PubMed=1750699;
RA Kassel D.B., Williams K.P., Musselman B.D., Smith J.A.;
RT "Optimization of the fragmentation in a fit-fast atom bombardment ion
RT source for the sequencing of peptides at the picomole level.";
RL Anal. Chem. 63:1978-1983(1991).
FT NON_TER 1
FT NON_TER 8
SQ SEQUENCE 8 AA; 772 MW; 783DDAADC2C732C8 CRC64;

Query Match 16.7%; Score 1; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 G 1
Db 1 G 1

RESULT 79
P82858
ID P82858 PRELIMINARY; PRT; 8 AA.
AC P82858;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Uricase (EC 1.7.3.3) (Urate oxidase) (Fragment).
OS Puccinia recondita f. sp. triseti.
OC Eukaryota; Fungi; Basidiomycota; Urediniomycetes; Urediniomycetidae;
OC Uredinales; Pucciniaceae; Puccinia.
OX NCBI_TaxID=142679;
RN [1]
RP SEQUENCE, FUNCTION, AND DEVELOPMENTAL STAGE.
RC TISSUE=Spore;
RA Aguilar M., Montalbini P., Pineda M.;
RL Submitted (NOV-2000) to the SWISS-PROT data bank.
CC -!- FUNCTION: INVOLVED IN HOST-PARASITE RELATIONSHIP BETWEEN PLANTS
CC AND FUNGI.
CC -!- CATALYTIC ACTIVITY: URATE + O(2) + H(2)O = 5-HYDROXYISOURATE +
CC H(2)O(2) (5-HYDROXYISOURATE DECOMPOSE TO FORM ALLANTOIN).
CC -!- SUBCELLULAR LOCATION: PEROXISOMAL (BY SIMILARITY).
CC -!- DEVELOPMENTAL STAGE: GERMINATION.
CC -!- SIMILARITY: BELONGS TO THE URICASE FAMILY.
DR InterPro: IPR002042; Uricase.
DR PROSITE: PS00366; URICASE; PARTIAL.
KW Oxidoreductase; Purine metabolism; Peroxisome.
FT NON_TER 8
FT NON_TER 8
SQ SEQUENCE 8 AA; 777 MW; 98C1ADD735B9D76D CRC64;

Query Match 16.7%; Score 1; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6
Db 4 S 4

RESULT 80
Q9BY55
ID Q9BY55 PRELIMINARY; PRT; 8 AA.
AC Q9BY55;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CAMP responsive element moderator (Fragment).
GN CREM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21082082; PubMed=11214319;
RA Murphy W.J., Elzirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA O'Brien S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001).
RL EMBL: AY011664; AAG47575.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 8 AA; 1006 MW; DF02C331EEAB572A CRC64;

Query Match 16.7%; Score 1; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 H 2
Db 5 H 5

RESULT 81
Q15889
ID Q15889 PRELIMINARY; PRT; 8 AA.
AC Q15889;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE (Clone XP15H88) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA Coolbaugh M.I., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,
RA Caskey C.T.H.;
RT "Isolation of chromosome-specific genes by reciprocal probing of
RT arrayed cDNAs and cosmid libraries.";
RL Hum. Mol. Genet. 0:0-0(1995).
DR EMBL: L32070; AAT3879.1; -.
FT NON_TER 1
FT NON_TER 8
SQ SEQUENCE 8 AA; 865 MW; 0474472325A761E7 CRC64;

Query Match 16.7%; Score 1; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 H 2
Db 2 H 2

RESULT 82
Q15900
ID Q15900 PRELIMINARY; PRT; 8 AA.
AC Q15900;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE (Clone XP7B11a) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA Coolbaugh M.I., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,
RA Caskey C.T.H.;

```


DR EMBL; AJ243097; CAB45408.1; -

FT NON TER 8 8
SQ SEQUENCE 8 AA; 767 MW; EE6EBDEB862D5B6 CRC64;

Query Match 16.7%; Score 1; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 S 6
DB 2 S 2

RESULT 87

ID Q15901 PRELIMINARY; PRT; 8 AA.
AC Q15901;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE (Clone XP7B11B) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA Coolbaugh M.I., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,
RA Caskey C.T.H.;
RT "Isolation of chromosome-specific genes by reciprocal probing of
RT arrayed cDNAs and cosmid libraries."
RL Hum. Mol. Genet. 0:0-0(1995).
DR EMBL; L32080; AAA73891.1; -.
FT NON TER 1 1
FT NON TER 8 8
SQ SEQUENCE 8 AA; 860 MW; 37D72878676729CB CRC64;

Query Match 16.7%; Score 1; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 G 1
DB 5 G 5

RESULT 88

ID Q9Y4J4 PRELIMINARY; PRT; 8 AA.
AC Q9Y4J4;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TReMBLrel. 12, Last annotation update)
DE Runt/82nt/MTG8 protein (Fragment).
GN Runt/82nt/MTG8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95002916; PubMed=7919324;
RA Tighe J.E., Calabi F.;
RT "Alternative, out-of-frame runt/MTG8 transcripts are encoded by the
RT derivative (8) chromosome in the t(8;21) of acute myeloid leukemia
RT M2."
RL Blood 84:2115-2121(1994).
DR EMBL; S74092; AAD14144.1; -.
FT NON TER 1 1
FT NON TER 8 AA; 1067 MW; 20F414044B17244B CRC64;
SQ SEQUENCE 8 AA; 1067 MW; 20F414044B17244B CRC64;

Query Match 16.7%; Score 1; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 H 2
DB 8 H 8

RESULT 89

ID Q15890 PRELIMINARY; PRT; 8 AA.
AC Q15890;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE (Clone XP19G12A) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA Coolbaugh M.I., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,
RA Caskey C.T.H.;
RT "Isolation of chromosome-specific genes by reciprocal probing of
RT arrayed cDNAs and cosmid libraries."
RL Hum. Mol. Genet. 0:0-0(1995).
DR EMBL; L32083; AAA73880.1; -.
FT NON TER 1 1
FT NON TER 8 8
SQ SEQUENCE 8 AA; 975 MW; 605EA6C5BEA5A2D3 CRC64;

Query Match 16.7%; Score 1; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 S 6
DB 3 S 3

RESULT 90

ID Q9UDZ4 PRELIMINARY; PRT; 8 AA.
AC Q9UDZ4;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE RHD protein (Fragment).
GN RHD
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=97260406; PubMed=9106526;
RA Matsessi G., Cherif-Zahar B., Mouro I., Cartron J.P.;
RT "Characterization of the recombination hot spot involved in the
RT genomic rearrangement leading to the hybrid D-CE-D gene in the DVI
RT phenotype."
RL Am. J. Hum. Genet. 60:808-817(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=93066356; PubMed=1438298;
RA Le Van Kim C., Mouro I., Cherif-Zahar B., Raynal V., Cherrier C.,
RA Cartron J.P., Colin Y.;
RT "Molecular cloning and primary structure of the human blood group Rh

T polypeptide.*;
Proc. Natl. Acad. Sci. U.S.A. 89:10925-10929(1992).
EMBL: Z97031; CAB09727.1; -;
FT NON_TER 1 1
SQ SEQUENCE 8 AA; 1042 MW; D296944691FBSAB1 CRC64;

Query Match 16.7%; Score 1; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 H 2
|
Db 4 H 4

RESULT 91
Q15898 PRELIMINARY; PRT; 8 AA.
AC Q15898;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE (Clone XP15H8A) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA Coolbaugh M.I., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,
RA Caskey C.T.H.;
RT "Isolation of chromosome-specific genes by reciprocal probing of
RT arrayed cDNAs and cosmid libraries.";
RL Hum. Mol. Genet. 0:0-0(1995).
DR EMBL; L32078; AAA73888.1; -;
FT NON_TER 1 1
SQ SEQUENCE 8 AA; 938 MW; 34A415B0477B45BB CRC64;

Query Match 16.7%; Score 1; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6
|
Db 2 S 2

RESULT 92
Q15893 PRELIMINARY; PRT; 8 AA.
AC Q15893;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE (Clone XP587A) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA Coolbaugh M.I., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,
RA Caskey C.T.H.;
RT "Isolation of chromosome-specific genes by reciprocal probing of
RT arrayed cDNAs and cosmid libraries.";
RL Hum. Mol. Genet. 0:0-0(1995).
DR EMBL; L32073; AAA73883.1; -;

FT NON_TER 1 1
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 874 MW; DAA1B6D7376456C5 CRC64;

Query Match 16.7%; Score 1; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6
|
Db 1 S 1

RESULT 93
Q15888 PRELIMINARY; PRT; 8 AA.
AC Q15888;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE (Clone XP15H8A) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA Coolbaugh M.I., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,
RA Caskey C.T.H.;
RT "Isolation of chromosome-specific genes by reciprocal probing of
RT arrayed cDNAs and cosmid libraries.";
RL Hum. Mol. Genet. 0:0-0(1995).
DR EMBL; L32069; AAA73878.1; -;
FT NON_TER 1 1
SQ SEQUENCE 8 AA; 1068 MW; 0315A37EAB5B0763 CRC64;

Query Match 16.7%; Score 1; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 S 6
|
Db 7 S 7

RESULT 94
Q16468 PRELIMINARY; PRT; 8 AA.
AC Q16468;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE DNA for cosmid CC13-1134 PCR primer 1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96435920; PubMed=8838806;
RA James L.A., Ogilvie D.J., Yamakawa K., Nakamura Y., Stirling C.J.,
RA Anand R.;
RT "Walking, cloning and mapping with YACs in 3q27. Localisation of 5
RT ESTs including 3 members of the Cystatin gene family and
RT identification of CpG islands.";
RL Genomics 32:425-430(1996).
DR EMBL; X88976; CAA61407.1; -;
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 925 MW; FDS411A7376871E6 CRC64;


```

Query Match      16.7%; Score 1; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 H 2
Db      2 H 2

RESULT 95
Q9UCN4
ID Q9UCN4 PRELIMINARY; PRT; 8 AA.
AC Q9UCN4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Cell-surface heparin/HEPARANSULFATE-binding protein peptide 3
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE
RX MEDLINE=92291065; PubMed=1601862;
RA Raboudi N., Julian J., Rohde L.H., Carlson D.D.;
RT "Identification of cell-surface heparin/heparan sulfate-binding
RT proteins of a human uterine epithelial cell line (RL95).";
RL J. Biol. Chem. 267:11930-11939(1992).
SQ SEQUENCE 8 AA; 689 MW; 8098733DD33DD87D CRC64;

Query Match      16.7%; Score 1; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 G 1
Db      2 G 2

RESULT 96
Q9UNC7
ID Q9UNC7 PRELIMINARY; PRT; 8 AA.
AC Q9UNC7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE SHMT protein (Fragment).
DE SHMT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Chave K.J., Snell K., Sanders P.G.;
RT "Isolation and characterization of human genomic sequences encoding
RT cytosolic serine hydroxymethyltransferase.";
RL Biochem. Soc. Trans. 25:53-53(1997).
DR EMBL: Y14492; CAB54844.1; -.
FT NON_TER 1
FT NON_TER 8
SQ SEQUENCE 8 AA; 868 MW; 7C205721E44AB5B8 CRC64;

Query Match      16.7%; Score 1; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 G 1
Db      1 G 1

RESULT 97
Q9HCQ0
ID Q9HCQ0 PRELIMINARY; PRT; 8 AA.
AC Q9HCQ0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Phosphodiesterase 10A7 (PDE10A7) (Fragment).
GN HSPDE10A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20453115; PubMed=10998054;
RA Fujishige K., Kotera J., Yuasa K., Omori K.;
RT "The human phosphodiesterase PDE10A gene. Genomic organization and
RT evolutionary relatedness with other PDEs containing GAP domains.";
RL Eur. J. Biochem. 267:5943-5951(2000).
DR EMBL: AB041779; BAB16368.1; -.
FT NON_TER 8
FT NON_TER 8
SQ SEQUENCE 8 AA; 966 MW; FD4B19D5A6C76446 CRC64;

Query Match      16.7%; Score 1; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      6 S 6
Db      5 S 5

RESULT 98
Q9P285
ID Q9P285 PRELIMINARY; PRT; 8 AA.
AC Q9P285;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Clotting factor VIII (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Shibata M., Shima M., Morichika S., Yoshiola A.;
RT "Human clotting factor VIII gene, junction regions of the deletion of
RT exon 4 through 7.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB040872; BAA94312.1; -.
FT NON_TER 1
FT NON_TER 8
SQ SEQUENCE 8 AA; 866 MW; 1C16987AAB05BDD3 CRC64;

Query Match      16.7%; Score 1; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      6 S 6
Db      3 S 3

RESULT 99
Q8TF70
ID Q8TF70 PRELIMINARY; PRT; 8 AA.
AC Q8TF70;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Microphthalmia-associated transcription factor (Fragment).

```

MITF.
 NS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 [1]
 RN SEQUENCE FROM N.A.
 RA Takeda K., Yasumoto K., Shibahara S.;
 RA "An Mitf isoform exclusively expressed in the affected cells of Mitf
 RT mutants.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 OR EMBL; AB061771; BAB85121.1; --
 OR NON TER 8
 PT SEQUENCE 8 AA; 989 MW; ED0727204415A1A6 CRC64;
 SEQ

Query Match 16.7%; Score 1; DB 4; Length 8;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

6 S 6
 3 S 3

RESULT 100
 Q9P0K3 PRELIMINARY; PRT; 8 AA.
 ID Q9P0K3
 AC Q9P0K3
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
 DE cGMP-specific phosphodiesterase PDE5A2 (Fragment).
 EN PDE5A.
 GN Homo sapiens (Human).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 [1]
 RN SEQUENCE FROM N.A.
 RA MEDLINE=20145478; PubMed=10673249;
 RA Lin C.S., Lau A., Tu R., Lue T.F.;
 RA "Identification of three alternative first exons and an intronic
 RT promoter of human PDE5A gene.";
 RL Biochem. Biophys. Res. Commun. 268:596-602(2000).
 OR EMBL; AF155195; AAP40302.1; --
 OR NON TER 8
 PT SEQUENCE 8 AA; 908 MW; E8A33AA879D76726 CRC64;
 SEQ

Query Match 16.7%; Score 1; DB 4; Length 8;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 G 1
 5 G 5

Search completed: October 21, 2003, 18:27:10
 Job time : 53.6667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:11:55 ; Search time 51.3333 Seconds
(without alignments)
18.552 Million cell updates/sec

Title: US-10-057-890A-15

Perfect score: 6

Sequence: 1 GHHS 6

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 251420

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Listing first 150 summaries

Database : A_Geneseq_19Jun03.*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5	83.3	6	13	AA231933
2	5	83.3	6	24	ABP97637
3	5	83.3	8	21	AA00161
4	5	83.3	9	17	AAW06142
5	5	83.3	9	17	AA00161
6	5	83.3	9	20	AA28306
7	5	83.3	9	21	AA44463
8	5	83.3	9	22	AAE12682
9	5	83.3	9	22	AAE2692
10	5	83.3	9	22	AAE2692
11	5	83.3	9	22	AAE2692
12	5	83.3	9	22	AAE2692
13	5	83.3	9	22	AAE2692
14	5	83.3	9	22	AAE2692
15	5	83.3	9	22	AAE2692
16	5	83.3	9	22	AAE2692
17	5	83.3	9	22	AAE2692
18	5	83.3	9	22	AAE2692
19	5	83.3	9	22	AAE2692
20	5	83.3	9	22	AAE2692
21	5	83.3	9	22	AAE2692
22	5	83.3	9	22	AAE2692
23	5	83.3	9	22	AAE2692
24	5	83.3	9	22	AAE2692
25	5	83.3	9	22	AAE2692
26	5	83.3	9	22	AAE2692
27	5	83.3	9	22	AAE2692
28	5	83.3	9	22	AAE2692
29	5	83.3	9	22	AAE2692
30	5	83.3	9	22	AAE2692
31	5	83.3	9	22	AAE2692
32	5	83.3	9	22	AAE2692
33	5	83.3	9	22	AAE2692
34	5	83.3	9	22	AAE2692
35	5	83.3	9	22	AAE2692
36	5	83.3	9	22	AAE2692
37	5	83.3	9	22	AAE2692
38	5	83.3	9	22	AAE2692
39	5	83.3	9	22	AAE2692
40	5	83.3	9	22	AAE2692
41	5	83.3	9	22	AAE2692
42	5	83.3	9	22	AAE2692
43	5	83.3	9	22	AAE2692
44	5	83.3	9	22	AAE2692
45	5	83.3	9	22	AAE2692
46	5	83.3	9	22	AAE2692
47	5	83.3	9	22	AAE2692
48	5	83.3	9	22	AAE2692
49	5	83.3	9	22	AAE2692
50	5	83.3	9	22	AAE2692
51	5	83.3	9	22	AAE2692
52	5	83.3	9	22	AAE2692
53	5	83.3	9	22	AAE2692
54	5	83.3	9	22	AAE2692
55	5	83.3	9	22	AAE2692
56	5	83.3	9	22	AAE2692
57	5	83.3	9	22	AAE2692
58	5	83.3	9	22	AAE2692
59	5	83.3	9	22	AAE2692
60	5	83.3	9	22	AAE2692
61	5	83.3	9	22	AAE2692
62	5	83.3	9	22	AAE2692
63	5	83.3	9	22	AAE2692
64	5	83.3	9	22	AAE2692
65	5	83.3	9	22	AAE2692
66	5	83.3	9	22	AAE2692
67	5	83.3	9	22	AAE2692
68	5	83.3	9	22	AAE2692
69	5	83.3	9	22	AAE2692
70	5	83.3	9	22	AAE2692
71	5	83.3	9	22	AAE2692
72	5	83.3	9	22	AAE2692
73	5	83.3	9	22	AAE2692
74	5	83.3	9	22	AAE2692
75	5	83.3	9	22	AAE2692
76	5	83.3	9	22	AAE2692
77	5	83.3	9	22	AAE2692
78	5	83.3	9	22	AAE2692
79	5	83.3	9	22	AAE2692
80	5	83.3	9	22	AAE2692
81	5	83.3	9	22	AAE2692
82	5	83.3	9	22	AAE2692

E coli outer membr
Molecular interact
Protein allergen l
Dermatophagoides s
N-terminal of pept
Leader sequence of
Human TRFP constu
His-Tag containing
Thioredoxin relate
Epitope tag #2 use
Hexa-histidine tag
9 Histidine peptid
O6-alkylguanine-DN
HIS tag peptide us
Multidomain protei
Isoelectric point
Chelating peptide
Chelating peptide
Atomic force micro
Transition metal i
Transition metal i
Isoelectric point
Chelating peptide
Tag #1, to facilit
Purification tag o
Isoelectric point
Isoelectric point
Hexa-histidine tag
Poly-His tag for c
Poly-His tag for c
Poly-histidine pep
Nickel ion binding
Single chain antib
VH-VL domain linke
Sequence of a supp
Epitope tag. Synt
Major outer membra
Human TAK1 6xHis p
Peptide used in th
peptide comprising
Histidine tag rela
B. subtilis lumazi
Polyhistidine tag
His6 tag used to c
6-His peptide epit
His-6 peptide SEQ
Hexahistidine pept
His tag. Unidenti
Peptide tag 1. Sy
Epitope tag #7 use
F5H alpha or beta
Hexahis tag used t
His tag used for r
Peptide which inhi
Poly-His tag, to g
6-His epitope tag
Amino acid sequenc
His tag useful as
Metal capturing pr
His-tag peptide.
6-His epitope used
C-terminal HIS tag
Poly-His tag used
His Tag for purifi
Nascent protein de
Peptide #7, a comp
6-His epitope. Un
Transdominant effe
6-His peptide epit
HIS tag peptide.

83 4 66.7 6 23 ABP56408 Thioredoxin relate
84 4 66.7 6 23 AAE28476 6HIS peptide used
85 4 66.7 6 23 AAE28606 His peptide tag.
86 4 66.7 6 23 AAO22927 Human papillomavir
87 4 66.7 6 23 ABG32856 His tag for yeast
88 4 66.7 6 23 ABG33013 Synthetic 6-His ta
89 4 66.7 6 23 AAE26107 His tag peptide us
90 4 66.7 6 23 ABG32021 Synthetic protein
91 4 66.7 6 23 AAE23795 His6 tag used in t
92 4 66.7 6 23 AAB77485 AAV-helper plasmid
93 4 66.7 6 23 AAM49702 Xpni/Sali His-tag
94 4 66.7 6 23 AAO19983 Histidine (His)-ta
95 4 66.7 6 23 AAU76910 C terminal His tag
96 4 66.7 6 23 AAE18829 His tag used in th
97 4 66.7 6 23 AAU75289 His6 tag used in c
98 4 66.7 6 23 AAE16561 Epitope tag #5 fus
99 4 66.7 6 23 AAU10567 Purification pepti
100 4 66.7 6 24 ABU08057 Synthetic Histidin
101 4 66.7 6 24 ABU08159 Synthetic poly His
102 4 66.7 6 24 AAE34362 His tag used to il
103 4 66.7 6 24 AAE34693 His tag peptide.
104 4 66.7 6 24 AAO16698 PAL-781 plasmid-re
105 4 66.7 6 24 ABUS8805 Mucin 1 (MUC-1) sp
106 4 66.7 6 24 ABP58380 Histidine tag. Sy
107 4 66.7 6 24 ABG73745 His-tag #1. Synth
108 4 66.7 6 24 ABUS5850 His tag peptide.
109 4 66.7 6 24 AAO16240 Synthetic peptide
110 4 66.7 6 24 AAE32846 His-tag peptide.
111 4 66.7 6 24 ABP56478 Interferon gamma v
112 4 66.7 6 24 AAO16087 Hexa-histidine pep
113 4 66.7 6 24 AAO16087 Neurological/CNS d
114 4 66.7 6 24 AAE30397 6-HIS tag used to
115 4 66.7 6 24 AAE30831 HIS tag peptide us
116 4 66.7 6 24 ABP55464 Peptide SEQ ID NO:
117 4 66.7 6 24 ABP55533 Adapter peptide SE
118 4 66.7 6 24 ABG98509 Synthetic peptide
119 4 66.7 7 9 AAP80401 Affinity peptide c
120 4 66.7 7 19 AAM60081 Mus musculus RH ep
121 4 66.7 7 20 AAY03173 Human cardiac trop
122 4 66.7 7 21 AAY95081 Histidine tail seq
123 4 66.7 7 21 AAY98190 Histidine tail seq
124 4 66.7 7 22 AAG64363 Peptide insert #3.
125 4 66.7 7 22 AAB31975 Antigen used to ra
126 4 66.7 7 23 AAE25397 His tag peptide.
127 4 66.7 7 23 AAE14511 Tag added to C-ter
128 4 66.7 8 16 AAR77412 Extension peptide
129 4 66.7 8 17 AAR95439 Polystyridine tail
130 4 66.7 8 18 AAM20036 C-terminal metal b
131 4 66.7 8 19 AAM61479 A. fumigatus aller
132 4 66.7 8 21 AAB09991 C. tetani tetanus
133 4 66.7 8 21 AAB03014 Tag fused to MN pr
134 4 66.7 8 21 AAY96685 Polystyridine affi
135 4 66.7 8 21 AAY53237 MN protein C-termi
136 4 66.7 8 22 AAM52174 Peptide tag 2. Sy
137 4 66.7 8 22 AAE13200 His tag used for a
138 4 66.7 8 22 AAM51721 FSH alpha or beta
139 4 66.7 8 22 AAE11985 Tag #2, to facilit
140 4 66.7 8 22 AAE08227 Tag #1 for human i
141 4 66.7 8 22 AAG64395 Peptide tag. Unid
142 4 66.7 8 22 AAG64364 Peptide insert #4.
143 4 66.7 8 22 AAU00164 His Tag (2) for pu
144 4 66.7 8 22 AAB48167 Histidine tag to c
145 4 66.7 8 22 AAB36993 Peptide used in re
146 4 66.7 8 23 ABG71814 Peptide tag for fa
147 4 66.7 8 23 ABG31736 Fish growth hormon
148 4 66.7 8 23 ABG32022 Synthetic protein
149 4 66.7 8 23 ABP53086 Crp 27C1 C-termina
150 4 66.7 8 23 AAU99299 Brevibacillus Chos

ALIGNMENTS

RESULT 1
AAR23933
ID AAR23933 standard; Protein; 6 AA.
XX AAR23933;
XX
DT 25-MAR-2003 (updated)
DT 15-NOV-1992 (first entry)
XX
DE Lactoferrin-introduced peptide (3).
XX Human lactoferrin; LF; transformation; ss.
XX Synthetic.
XX JF04108390-A.
XX 09-APR-1992.
XX 29-AUG-1990; 90JP-0227619.
XX 29-AUG-1990; 90JP-0227619.
XX (SNOW) SNOW BRAND MILK PROD CO LTD.
XX WPI; 1992-171655/21.
XX Introduction of foreign genes into animal cells - using a
PT complex comprising the gene with lactoferrin.
XX
XX Disclosure; Page 5; 10pp; Japanese.
XX The sequences given in AAR23931 - AAR23938 are synthetic peptides which
CC were introduced into human lactoferrin (LF) by introducing the DNA
CC encoding the peptide into the human LF gene and then incubating the
CC mixture. The introduction of the peptide caused increased
CC transformation rates compared to unmodified LF and this method could
CC be used to introduce an exogenous gene into an animal cell with no
CC stress.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 6 AA;
Query Match 83.3%; Score 5; DB 13; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GHHHH 5
DB 1 GHHHH 5

RESULT 2
ABP97637
ID ABP97637 standard; peptide; 6 AA.
XX
AC ABP97637;
XX
DT 16-MAY-2003 (first entry)
XX
DE Proteolytic product of a NAD+-dependent DNA ligase.
XX NAD+-dependent DNA ligase; enzyme; mutant; DNA ligation; domain Ia.
XX Amsacta moorei entomopoxvirus.
XX WO2003000858-A2.
XX
XX 03-JAN-2003.
XX
XX 24-JUN-2002; 2002WO-US20005.
XX 24-JUN-2001; 2001US-300727P.
PR

XX (SLOK) SLOAN KETTERING INST CANCER RES.
 XX Shuman S, Sriskanda V;
 XX WPI; 2003-221413/21.
 DR
 XX Novel NAD-dependent DNA ligase, defective in its reaction with NAD but
 XX active in the ligation of pre-adenylated DNA nicks, useful for
 XX screening compounds, comprises a mutated domain Ia of DNA ligase
 XX
 XX Example 4; Fig 7; 92pp; English.
 PS
 XX Peptides ABP97634-39 represent the proteolytic products of a
 CC NAD+-dependent DNA ligase. The specification describes a NAD+-dependent
 CC DNA ligase enzyme that is defective in its reaction with NAD+ but is
 CC active in the ligation of pre-adenylated DNA nicks. This enzyme is
 CC mutated by deleting domain Ia of the DNA ligase or by substituting one
 CC or more conserved residues in the domain Ia. The mutant enzyme is useful
 CC for screening for a compound that binds to the NAD+ substrate recognition
 CC site of NAD+-dependent DNA ligase. Screening methods using the mutant
 CC enzyme are geared to detect specific candidate drugs that interfere with
 CC NAD+ recognition by ligases based on biochemically validated differences
 CC in binding to the ligases. It is also useful for selecting and
 CC identifying small molecule ligands that specifically bind to the NAD+
 CC site.
 XX
 XX Sequence 6 AA;
 SQ
 Query Match 83.3%; Score 5; DB 24; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GHHHH 5
 DB 1 GHHHH 5
 |||||
 |||||
 RESULT 3
 AAB00161
 ID AAB00161 standard; Peptide; 8 AA.
 XX
 AC AAB00161;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Hexahistidine tag used in sCD4-SCFv(17b) fusion protein.
 XX
 KW Fusion protein; HIV; human immunodeficiency virus; antibody; Fv;
 KW AIDS; acquired immune deficiency syndrome; neutralisation;
 KW infection; gene therapy; CD4; gp120; glycoprotein; resistance;
 KW vaccination; binding domain; single chain antibody; chimera;
 KW chimeric protein.
 XX
 OS Synthetic.
 XX
 WO200055207-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 16-MAR-2000; 2000WO-US06946.
 XX
 PR 16-MAR-1999; 99US-0124681.
 XX
 PA (USSH) US NAT INST OF HEALTH.
 XX
 PI Berger EA, Del Castillo CM;
 XX
 DR WPI; 2000-638183/61.
 XX
 XX Novel neutralizing bispecific fusion proteins effective in viral such
 XX as HIV neutralization, comprises two different binding domains,
 XX inducing-binding domain and induced-binding domain functionally linked

PT by linker
 XX
 PS Example 1; Page 49; 55pp; English.
 XX
 CC sCD4-SCFv(17b) is a neutralising bispecific fusion protein capable of
 CC binding to two sites of its target protein. The protein comprises a
 CC first binding domain capable of binding to an inducing site on the
 CC target protein, a second binding domain capable of forming
 CC neutralising complex with an induced epitope of the target protein and
 CC a linker connecting the binding domains. sCD4-SCFv(17b) comprises a
 CC soluble CD4 fragment (containing domains D1 and D2) fused to a single
 CC chain Fv portion of antibody 17b via a linker. sCD4-SCFv(17b), its
 CC variant, analogue or mimetic is used for inactivating gp120 protein of
 CC HIV, and for neutralising HIV. It is also used for blocking and
 CC preventing the binding of the viral or recombinant gp120 protein to
 CC soluble CD4 or lymphocyte CD4 and for inhibiting HIV replication. The
 CC chimeric proteins is therefore useful for treating HIV infection and
 CC also AIDS. It is particularly useful in the prevention of
 CC infection during or immediately after HIV exposure (e.g.,
 CC mother/infant transmission, post-exposure prophylaxis, and as a
 CC topical inhibitor) and for providing long term resistance to HIV
 CC infections and AIDS. Gene therapy is used to secrete the bispecific
 CC protein at mucosal surfaces, such as the vaginal, rectal or oral
 CC mucosa. The fusion proteins is highly potent, broadly cross-reactive
 CC with neutralising antibody with high in vivo activity and no
 CC FC-mediated undesirable targeting properties. When the fusion
 CC protein is substantially derived from human proteins, it has minimal
 CC immunogenicity and toxicity in humans which is of great value in
 CC prevention of infection during or immediately after HIV exposure.
 CC To generate the fusion protein the full length CD4 coding sequence
 CC was excised from the plasmid pCB-3 using restriction enzymes StuI
 CC and SpeI. Synthetic oligonucleotides (AAA54049, AAA54046) were ligated
 CC and used to regenerate the cleaved StuI site and the next two bases
 CC of the CD4 cDNA and to produce an SpeI overhang at the 3' end. The
 CC sequence also reconstructs the second domain of CD4 and encodes the
 CC 37 amino acid intermediate linker (Gly_4Ser)_6Gly_4Thr_2Ser followed
 CC directly by the universal translation termination signal. A BamHI
 CC site is included near the end of the third (Gly_4Ser) repeat to
 CC enable linkage to the 17b SCFv with the exact linker sequence. The
 CC resulting intermediate plasmid was designated pCD1 but the sCD4
 CC linker insert was then subcloned into a pSC59 derivative lacking a
 CC BamHI site, forming pCD2. The 17b SCFv sequence was amplified from
 CC its starting plasmid vector pmt del 0 using two primers (AAA54047,
 CC AAA54048). The amplified fragment was digested with BamHI and SpeI
 CC and force-cloned into pCD2 to generate the sCD4-SCFv(17b) fusion
 CC protein in a plasmid designated pCD3.
 XX
 XX Sequence 8 AA;
 SQ
 Query Match 83.3%; Score 5; DB 21; Length 8;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GHHHH 5
 DB 2 GHHHH 6
 |||||
 |||||
 RESULT 4
 AAWC6142
 ID AAWC6142 standard; peptide; 9 AA.
 XX
 AC AAWC6142;
 XX
 DT 25-MAR-2003 (updated)
 DT 03-FEB-1997 (first entry)
 XX
 DE Variant adenovirus proteinase cleavable carboxy-terminal peptide.
 XX
 KW Adenovirus proteinase; polyanion cofactor; peptide cofactor;
 KW antibody; detection; purification; antiviral; vaccine; treatment;
 KW cleavage.
 XX

```

S      Synthetic.
X      US5543264-A.
N      06-AUG-1996.
XX     19-NOV-1993; 93US-0155171.
XX     19-NOV-1993; 93US-0155171.
XX     29-JUN-1990; 90US-0545585.
XX     13-MAR-1992; 92US-0851217.
XX     (ASUY-) ASSOC UNIVERSITIES INC.
XX     Anderson CW, Mangel WF;
XX     WPI; 1996-370630/37.
XX     Recombinant adenovirus proteinase activatable with polyanion
XX     co-factor - used to develop prods. for use in the detection,
XX     treatment and prevention of adenovirus infection
XX     Claim 15; Column 80; 68pp; English.
XX     A recombinant activatable adenovirus proteinase (AP) and variants of
XX     adenovirus type 2 proteinase (A2P) which has an additional cleavable
XX     carboxy-terminal peptide of about 9 amino acids are new. Also new
XX     is a reconstituted preparation of activated AP which comprises AP; a
XX     polyanion cofactor of AP activity and a peptide cofactor of AP
XX     activity. The products can be used for producing large amounts of
XX     activated AP's. The AP's can be used for the specific cleavage of
XX     proteins e.g. to purify proteins or portions of proteins or for
XX     deletion of protein domains. The products can also be used to
XX     produce AP's with altered activities and to identify substances
XX     which affect AP activity which can be used for treating adenovirus
XX     infections or modulating adenovirus vaccines. The AP's can also be
XX     used for producing antibodies which can be used for detection,
XX     purification and study of adenoviruses. Antibodies to the peptide
XX     cofactor can also be produced which can be used for studying the
XX     function of the cofactor or as antiviral agents adenovirus. This
XX     sequence represents the additional cleavable carboxy-terminal
XX     peptide mentioned above.
XX     (Updated on 25-MAR-2003 to correct PF field.)
XX     Sequence 9 AA;
XX     Query Match 83.3%; Score 5; DB 17; Length 9;
XX     Best Local Similarity 100.0%; Pred. No. 9.3e+05;
XX     Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY     1 GHHHH 5
DB     1 GHHHH 5
XX     RESULT 5
XX     IID AAR97378 standard; Peptide; 9 AA.
XX     AC AAR97378;
XX     03-OCT-1996 (first entry)
XX     DE Penta-histidine peptide spacer.
XX     KW Antigen-binding fusion protein; antibody engineering;
XX     immunoeffector; cytostatic; phospholipase A activating protein;
XX     PLAP; tumour necrosis factor; TNF; cancer; therapy.
XX     OS Synthetic.
XX     PN WO9611955-A1.
XX
PD     25-APR-1996.
XX     13-OCT-1995; 95WO-US12840.
XX     13-OCT-1994; 94US-0323445.
XX     (ENZO-) ENZON INC.
XX     Filpula D, Shorr R, Whitlow M;
XX     WPI; 1996-221949/22.
XX     New antigen binding proteins contg. immunoeffector or cytolytic
XX     peptide - attached to variable regions of antibody light or heavy
XX     chain, useful in diagnosis and in destroying cancer cells
XX     Claim 16; Page 49; 72pp; English.
XX     A peptide spacer (AAR97378) used in novel antigen binding proteins
XX     contains a penta-histidine portion that facilitates affinity
XX     purification. The spacer is used to link an immunoeffector
XX     or cytolytic peptide to the variable regions of antibody light
XX     or heavy chains. It is pref. inserted between a serine residue
XX     of CC49 VH and a glutamate residue of phospholipase A activating
XX     protein immunoeffector (see also (AAR97368). Single-chain or
XX     multivalent constructs can be produced (see also AAR97388-89)
XX     that are useful for cancer diagnosis or treatment.
XX     Sequence 9 AA;
XX     Query Match 83.3%; Score 5; DB 17; Length 9;
XX     Best Local Similarity 100.0%; Pred. No. 9.3e+05;
XX     Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY     2 HHHHS 6
DB     4 HHHHS 8
XX     RESULT 6
XX     IID AAY28306 standard; Protein; 9 AA.
XX     AC AAY28306;
XX     20-MAR-2003 (updated)
XX     28-SEP-1999 (first entry)
XX     DE Amino acid sequence of the wildtype Ad2 proteinase extension.
XX     KW adenovirus; proteinase; genomic; amino acid; Ad2.
XX     OS Adenovirus subgroup C.
XX     PN US5935840-A.
XX     10-AUG-1999.
XX     05-MAY-1995; 95US-0435998.
XX     19-NOV-1993; 93US-0155171.
XX     29-JUN-1990; 90US-0545585.
XX     13-MAR-1992; 92US-0851217.
XX     05-MAY-1995; 95US-0435998.
XX     (ASUY-) ASSOC UNIVERSITIES INC.
XX     PA Anderson CW, Mangel WF;
XX     PI WPI; 1999-457604/38.
XX     Activated recombinant adenovirus proteinases

```

PS Disclosure; Column 45; 66pp; English.

XX The amino acid sequence of the wildtype Ad2 proteinase extension,
 CC from the subgroup C Adenovirus.
 CC Activated adenovirus proteinase preparations can be used, within
 CC a purified system or within a cellular extract, to cleave a selected
 CC protein at a specific site.
 CC The timing of the cleavage event is able to be controlled by
 CC combining the substrate with the proteinase and DNA co factor, and
 CC then initiating the cleavage by adding the peptide co factor.
 CC (Updated on 20-MAR-2003 to correct PF field.)

XX Sequence 9 AA;

Query Match 83.3%; Score 5; DB 20; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHHH 5
 |||||
 Db 1 GHHHH 5

RESULT 7

AA44463
 ID AAY44463 standard; peptide; 9 AA.

XX AAY44463;

XX 27-MAR-2000 (first entry)

XX Gly(His)8 tag, ghhhhhhh.

XX Interleukin; IL-17B; PRO1031; PRO1122; cytokine IL-17; Gly(His)8 tag;
 KW Fc/His fusion protein; probe; degenerative cartilaginous disorder.

XX Synthetic.

XX WO9960127-A2.

XX 25-NOV-1999.

XX 14-MAY-1999; 99WO-US10733.

XX 15-MAY-1998; 98US-0085579.

XX 23-DEC-1998; 98US-0113621.

XX (GETH) GENENTECH INC.

XX Chen J, Filvaroff E, Goddard A, Gurney AL, Li H, Wood WI;
 PI WPI; 2000-116314/10.

XX New polypeptides designated PRO1031 and PRO1122 used to treat a
 PT degenerative cartilaginous disorder -
 XX Example 9; Page 82; 141pp; English.

XX The present sequence is the Gly(His)8 tag, that is used to link to the
 CC C-terminal end of the PRO1031 and PRO1122 polypeptide. It is used to
 CC generate His fusion proteins.

XX Sequence 9 AA;
 Query Match 83.3%; Score 5; DB 21; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHHH 5
 |||||
 Db 1 GHHHH 5

RESULT 8

AAE12682
 ID AAE12682 standard; peptide; 9 AA.

XX AAE12682;

XX 03-JAN-2002 (first entry)

XX Polyhistidine tag peptide used in the expression of human ERbeta LBD.

XX Polyhistidine tag; multiplexed assay; therapeutic; diagnostic agent;
 KW molecular interaction; first member; second member; binding pair;
 KW human estrogen receptor beta ligand binding domain; ERbeta LBD.

XX Synthetic.

XX WO200175443-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US10506.

XX 31-MAR-2000; 2000US-193826P.

XX (GLAX) GLAXO GROUP LTD.

XX Consler TG, Gray JG, Iannone MA, Stimmel JB;
 PI WPI; 2001-639382/73.

XX Multiplex assay for determining relative binding of partners, useful
 PT for identifying diagnostic or therapeutic agents, by simultaneous
 PT reaction of individually labeled components -
 XX Example 3; Page 26; 73pp; English.

XX The present invention relates to a method for identifying, in a single
 CC assay, relative binding of two or more first members of a binding pair
 CC to one or more second members. The invention also provides a multiplexed
 CC assay for analysing complex molecular interactions. The method is used
 CC to identify potential therapeutic and diagnostic agents, including those
 CC that affect relative binding between the first and second members of a
 CC binding pair. The method is rapid, sensitive and suitable for
 CC multiplexed, high-throughput screening. The present sequence is a
 CC Polyhistidine tag peptide used in the expression of human estrogen
 CC receptor beta ligand binding domain (ERbeta LBD).

XX Sequence 9 AA;

Query Match 83.3%; Score 5; DB 22; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHHH 5
 |||||
 Db 4 GHHHH 8

RESULT 9

AA82692
 ID AAB82692 standard; Peptide; 9 AA.

XX AAB82692;

XX 15-OCT-2001 (first entry)

XX His-tag.

XX Acetylcholine binding protein; AChBP; mollusc;
 KW ligand-binding protein; ligand-gated ion channel; crystal;
 KW drug design; protein co-ordinate data; giant pond snail;
 KW schizophrenia; Alzheimer's disease; nicotine addiction;
 KW Tourette's syndrome; therapy; neurotropic; neuroprotective.

XX OS Synthetic.
 XX PN WO200158951-A2.
 XX PD 16-AUG-2001.
 XX PF 09-FEB-2001; 2001WO-EF01457.
 XX PR 10-FEB-2000; 2000EP-0200443.
 XX PR 31-OCT-2000; 2000EP-0203810.
 XX PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX PI Smit AB, Sixma TK;
 XX WPI; 2001-497071/54.
 XX PT Water-soluble ligand-binding proteins derived from molluscs and analogues
 XX PT of ligand-gated ion channels, useful in drug screening assay, where the
 XX PT drugs identified can be used in the treatment of Alzheimer's disease or
 XX PT schizophrenia -
 XX PS Example 4; Page 52; 260pp; English.
 XX CC The present sequence comprises a hexahistidine tag peptide,
 XX CC produced C-terminally to a recombinant acetylcholine binding
 XX CC protein (ACHBP) of the invention, specifically L-AChBP-T1 (see
 XX CC AAB82686) of the giant pond snail (*Lymnaea stagnalis*), upon
 XX CC production of the protein in a *Pichia pastoris* expression system.
 XX CC The invention relates to water-soluble ligand-binding proteins,
 XX CC based on ACHBPs and derived from molluscs, and analogues of
 XX CC ligand-gated ion channels, their crystals, and their use for
 XX CC screening ligands of ligand-gated ion channels. New drugs can be
 XX CC developed that selectively intervene in neuronal signalling
 XX CC pathways, especially where the ligand-gated ion channel is the
 XX CC nicotinic acetylcholine receptor, and the related disorder is
 XX CC Tourette's syndrome, Alzheimer's disease, addiction to nicotine or
 XX CC schizophrenia.
 XX CC
 XX SQ Sequence 9 AA;
 Query Match 83.3%; Score 5; DB 22; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GHHHH 5
 DB 3 GHHHH 7
 RESULT 10
 AA018574
 ID AA018574 standard; Peptide; 9 AA.
 XX AC
 XX AC AA018574;
 XX DT 15-OCT-2002 (first entry)
 XX DE E coli outer membrane protein OmpA modifying N-terminal tag.
 XX KW Outer membrane protein; OmpF; biosensor; membrane-spanning protein;
 XX KW protein array; tag.
 XX OS Synthetic.
 XX PN WO200257780-A1.
 XX PD 25-JUL-2002.
 XX PF 18-JAN-2002; 2002WO-GB00222.
 XX PR 18-JAN-2001; 2001GB-0001279.

PR 10-APR-2001; 2001GB-0008947.
 XX (UYNE-) UNIV NEWCASTLE VENTURES LTD.
 XX PI Lakey JH;
 XX WPI; 2002-608388/65.
 XX PT Novel product useful for detecting molecules e.g. polypeptides,
 XX PT comprises membrane-spanning protein coupled to a substrate, and lipid
 XX PT membrane formed from amphiphilic molecules and membrane-spanning
 XX PT protein molecules -
 XX PS Claim 33; Page 35; 52pp; English.
 XX CC The present invention relates to a product comprising a membrane-spanning
 XX CC protein, a lipid membrane formed from amphiphilic molecules and
 XX CC membrane-spanning protein molecules, and a substrate (the membrane
 XX CC protein is directly coupled to the substrate). This can be used for
 XX CC interaction with a molecule selected from polypeptide, antigenic
 XX CC polypeptides, antibodies or fragments of antibodies, receptors, ligands,
 XX CC antibiotics, drugs, pesticides, sugars, amino acids, fatty acids,
 XX CC peptides, hormones, steroids, nucleic acids, peptide nucleic acids,
 XX CC metals, and inorganic ions, particularly in a biosensor or protein array.
 XX CC The present sequence is a N-terminal tag used to modify the E. coli outer
 XX CC membrane protein OmpA.
 XX SQ Sequence 9 AA;
 Query Match 83.3%; Score 5; DB 23; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 HHHHS 6
 DB 4 HHHHS 8
 RESULT 11
 AAU84236
 ID AAU84236 standard; Peptide; 9 AA.
 XX AC
 XX AC AAU84236;
 XX DT 08-MAY-2002 (first entry)
 XX DE Molecular interaction assay associated polynhistidine tag.
 XX KW Molecular interaction assay; multiplexed assay;
 XX KW polynhistidine tag.
 XX OS Synthetic.
 XX PN US2002C04205-A1.
 XX PD 10-JAN-2002.
 XX PF 30-MAR-2001; 2001US-0821984.
 XX PR 30-MAR-2000; 2000US-193286P.
 XX PA (CONS/) CONSLER T G.
 XX PA (GRAY/) GRAY J G.
 XX PA (IANN/) IANNONE M A.
 XX PA (STIM/) STIMMEL J B.
 XX PI Consler TG, Gray JG, Iannone MA, Stimmel JB;
 XX WPI; 2002-163724/21.
 XX PT Analysing complex molecular interactions, by contacting members of
 XX PT binding pair to form the binding pair, coupling the members with

CC AAY50546-Y50555 represent house dust mite allergen peptide fragments
 CC derived from Der p 1, Der f II, Der f I and Der f II.
 XX
 SQ Sequence 10 AA;

Query Match 83.3%; Score 5; DB 20; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.68;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHH 5
 DB 2 GHHH 6

RESULT 14
 AAB28976
 ID AAB28976 standard; Peptide; 10 AA.

XX AAB28976;

XX 29-JAN-2001 (first entry)

XX N-terminal of peptides XYZ YXZ ZXY.

XX Cat; allergy; human T cell reactive feline protein; hTRFP;
 immunotherapy.

XX Pelis sp.

OS Synthetic.

XX US6120769-A.

XX 19-SEP-2000.

XX 28-APR-1995; 95US-0431184.

XX 02-SEP-1994; 94US-0300928.

XX 03-NOV-1989; 89US-0431565.

XX 28-FEB-1991; 91US-0662276.

XX 13-DEC-1991; 91US-0807529.

XX 25-MAR-1992; 92US-0857311.

XX 15-MAY-1992; 92US-0884718.

XX 15-JAN-1993; 93US-0006116.

XX (IMMU-) IMMULOGIC PHARM CORP.

XX Gefter ML, Garman RD, Greenstein JL, Bond JF;

XX WPI; 2000-601477/57.

XX Detecting, preventing and treating sensitivity to cat protein allergen
 PT comprises combining a biological sample with a human T cell reactive
 PT feline protein and determining the extent of binding that occurs -
 XX Example 25; Column 117; 106pp; English.
 XX The present invention relates to the detection of sensitivity to a cat
 XX protein allergen by combining a blood sample from a subject with a
 XX peptide of human T cell reactive feline protein (hTRFP). This method
 XX and the hTRFP peptides are useful for diagnosing, preventing and
 XX treating cat allergies by reducing or abolishing an individual's
 XX allergic response to a cat allergen. DNA encoding the TRFP may be
 XX used as probes to locate equivalent sequences present in other species.
 XX These may further be used to study the mechanism of immunotherapy of
 XX cat allergy, and to design modified derivatives, analogues or
 XX functional equivalents useful in immunotherapy. The present
 XX sequence was used in the invention.

XX Sequence 10 AA;

Query Match 83.3%; Score 5; DB 21; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.68;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHH 5
 DB 2 GHHH 6

RESULT 15

AAY87721
 ID AAY87721 standard; Protein; 10 AA.

XX AAY87721;

XX 22-AUG-2000 (first entry)

XX Leader sequence of peptides XYZ, YXZ and ZXY.

XX T-cell reactive feline protein; TRFP; Fel d I; cat allergen; primer;
 KW antiallergic; T cell stimulator; diagnostic; immunotherapy.

XX Pelis sp.

XX US6048962-A.

XX 11-APR-2000.

XX 27-APR-1995; 95US-0430014.

XX 02-SEP-1994; 94US-0300928.

XX 03-NOV-1989; 89US-0431565.

XX 28-FEB-1991; 91US-0662276.

XX 13-DEC-1991; 91US-0807529.

XX 25-MAR-1992; 92US-0857311.

XX 15-MAY-1992; 92US-0884718.

XX 15-JAN-1993; 93US-0006116.

XX (IMMU-) IMMULOGIC PHARM CORP.

XX Kuo M, Rogers BL, Gefter ML, Morgenstern JP, Brauer AW;

XX Greenstein JL, Griffith IJ, Garman RD;

XX WPI; 2000-316905/27.

XX New human T cell reactive feline protein useful for reducing or
 PT abolishing individual's allergic response to cat allergen comprising
 PT two different covalently linked peptide chains -

XX Example 25; Column 117-118; 106pp; English.

XX This invention describes a novel naturally occurring cat protein allergen
 CC (1), human T cell reactive feline protein (TRFP), comprising two
 CC different covalently linked peptide chains with a molecular weight of 20
 CC kD, 40 kD or 130 kD under non-reducing conditions and 5 kD or 10-18 kD
 CC under reducing conditions. The products of the invention have
 CC antiallergic activity and act as human T cell stimulators. TRFP is useful
 CC for reducing or preventing the adverse effects of cat allergens on cat
 CC allergic individuals and in ex vivo diagnostic tests to determine which
 CC peptides cause sensitivity so as to selectively use them to desensitize
 CC a cat sensitive individual. Purified TRFP is also useful for studying
 CC the mechanism of immunotherapy of cat allergy and to design modified
 CC derivatives, analogs or functional equivalents that are more useful in
 CC immunotherapy against cat allergy. DNA sequences encoding TRFP are
 CC useful as probes to locate equivalent sequences present in other species
 CC (goats, sheep, dogs, rabbits or horses) that may be useful in diagnostics
 CC and/or therapeutics. Fully defined and characterized TRFP provides
 CC complete and a very simple desensitization therapy. This sequence
 CC represents the leader sequence of peptides XYZ, YXZ and ZXY which are
 CC used in the method of the invention.

XX Sequence 10 AA;

Query Match 83.3%; Score 5; DB 21; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.68;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHHH 5
Db 2 GHHHH 6

RESULT 16
AAAY90141
ID AAAY90141 standard; Peptide; 10 AA.
XX AC
XX AC AAAY90141;
XX DT 13-JUL-2000 (first entry)
XX DE Leader sequence peptide.
XX KW Cat; TRFP; human T-cell reactive feline protein; cat protein allergen;
XX KW house dust; Fel d 1; cat allergy; Felis domesticus sensitivity; therapy;
XX KW diagnosis; goat; sheep; horse; rabbit; dog; leader sequence.
XX OS Unidentified.
XX PN US6025162-A.
XX XX
XX PD 15-FEB-2000.
XX PF 28-APR-1995; 95US-0430944.
XX PR 02-SEP-1994; 94US-0300928.
XX PR 03-NOV-1989; 89US-0431565.
XX PR 28-FEB-1991; 91US-0662276.
XX PR 13-DEC-1991; 91US-0807529.
XX PR 25-MAR-1992; 92US-0857311.
XX PR 15-MAY-1992; 92US-0884718.
XX PR 15-JAN-1993; 93US-0006116.
XX PA (IMMU-) IMMULOGIC PHARM CORP.
XX PI Morgenstern JP, Griffith IJ, Rogers BL;
XX WPI; 2000-181812/16.
XX DR
XX XX
XX PT New human T cell reactive feline protein, useful for desensitizing cat
XX PT allergic individuals to cat allergens -
XX PS Example 25; Column 60; 108pp; English.
XX CC This sequence represents a leader sequence peptide. The invention
XX CC relates to a human T cell reactive feline protein (TRFP). The protein is
XX CC a cat protein allergen, and was isolated from a vacuum bag extract
XX CC obtained by affinity purification of house dust collected from several
XX CC homes with cats. TRFP is composed of two covalently linked peptide
XX CC chains, and is also referred to as Fel d 1. TRFP and its peptides are
XX CC useful for reducing or preventing the adverse effects that exposure to
XX CC cat allergens normally has on cat allergic individuals (i.e. to
XX CC desensitize individuals to cat allergens or block the effect of the
XX CC allergens). TRFP is also used in methods of diagnosing sensitivity to
XX CC Felis domesticus in an individual. DNA sequences encoding TRFP can be
XX CC used as probes to locate equivalent sequences present in other species,
XX CC e.g. goat, sheep, horse, rabbit and dog, that may be useful in diagnostic
XX CC and/or therapeutic applications.
XX SQ Sequence 10 AA;
Query Match 83.3%; Score 5; DB 21; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.68;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHHH 5
Db 2 GHHHH 6

RESULT 18
AAAU19104
ID AAU19104 standard; Peptide; 10 AA.
XX AC
XX AC AAU19104;
XX DT 04-DEC-2001 (first entry)

RESULT 17
AAAY51518
ID AAAY51518 standard; Protein; 10 AA.
XX AC
XX AC AAAY51518;
XX DT 22-MAY-2000 (first entry)
XX DE Human TRFP construct leader sequence peptide.
XX KW T-cell reactive feline protein; TRFP; T cell epitope; T cell receptor;
XX KW down regulation; immune response; allergen; immunoglobulin E;
XX KW sensitivity; cat protein allergen; human.
XX OS Homo sapiens.
XX PN US6019972-A.
XX PD 01-FEB-2000.
XX PF 02-SEP-1994; 94US-0300928.
XX PR 03-NOV-1989; 89US-0431565.
XX PR 28-FEB-1991; 91US-0662276.
XX PR 13-DEC-1991; 91US-0807529.
XX PR 25-MAR-1992; 92US-0857311.
XX PR 15-MAY-1992; 92US-0884718.
XX PR 15-JAN-1993; 93US-0006116.
XX PA (IMMU-) IMMULOGIC PHARM CORP.
XX PI Garman RD, Greenstein JL, Kuo M, Briner TJ, Morville M, Gefter ML;
XX WPI; 2000-146862/13.
XX DR
XX PT Peptides of human T cell reactive feline protein for treating
XX PT sensitivity to cat protein allergens comprise at least one T cell
XX PT epitope recognized by a T cell receptor specific for the human T cell
XX PT reactive feline protein -
XX PS Example 25; Column 117-118; 105pp; English.
XX CC This invention describes a novel peptide (I) of human T cell reactive
XX CC feline protein (hTRFP) having at least one T cell epitope recognized
XX CC by a T cell receptor specific for the human T cell reactive feline
XX CC protein, the peptide consisting of at least 7-30 amino acids, and having
XX CC an amino acid sequence derived from an amino acid sequence comprising 94,
XX CC 96, 97, 109, or 111 residues, given in the specification. The peptides
XX CC down regulate the immune response to the allergen. The peptides have
XX CC reduced immunoglobulin E binding and reduce T cell responsiveness. The
XX CC peptide (I) is useful in compositions for treating sensitivity to a cat
XX CC protein allergen in a subject. This sequence represents the leader
XX CC sequence peptide used in the construction of peptides comprising of two
XX CC or more regions of human TRFP.
XX SQ Sequence 10 AA;
Query Match 83.3%; Score 5; DB 21; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.68;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHHH 5
Db 2 GHHHH 6

RESULT 18
AAAU19104
ID AAU19104 standard; Peptide; 10 AA.
XX AC
XX AC AAU19104;
XX DT 04-DEC-2001 (first entry)

D*	12-MAR-2003	(first entry)
XX		
DE		"Thioredoxin related peptide #3.
XX		
KW		Thioredoxin; TrxA; fusion protein; allogenic protein..
XX		
OS		Synthetic.
XX		
PN	CN189539-A.	
XX		
PD	05-AUG--1998.	
XX		
PF	31-JAN-1997;	97CN-0100362.
XX		
PR	31-JAN-1997;	97CN-0100362.
XX		
PA	(BIOL-) BIOLOGICAL ENG INST ACAD MILITARY MEDICI.	
PX		
PI	Cui L, Ma Q;	
XX		
DR	WPI; 2002-733391/80.	
XX		
PT	Soluble expression carrier and its fusion protein of novel thioredoxin	
PT	.	
XX		
PS	Disclosure; Page 3 (Disclosure); 14pp; Chinese.	
XX		
CC	The present invention describes a DNA molecule of a fusion protein which	
CC	is composed of the DNA sequence similar to encoded thioredoxin in	
CC	structure and the DNA sequence of an encoded allogenic protein. The	
CC	soluble expression of the fusion protein is efficient. The expressed	
CC	fusion protein can be easily purified as several amino acid sequences	
CC	are inserted at active sites of the thioredoxin gene sequence, and	
CC	hydrolysed by a specific protease as the junction between the two	
CC	encoded DNA sequences can be cut by the specific protease, obtaining a	
CC	recombined allogenic protein the same as the natural protein in one-class	
CC	structure. The present sequence represents a peptide which is given in	
CC	the exemplification of the present invention.	
XX		
SQ	Sequence 10 AA;	
	Query Match	83.3%; Score 5; DB 23; Length 10;
	Best Local Similarity	100.0%; Pred.No.0.68;
	Matches 5; Conservative	0; Mismatches 0; Indels 0; Gaps
Qy	1 GHHHH 5	
Db	2 GHHHH 6	
RESULT 20		
AAU97726		
ID	AAU97726 standard; Peptide; 10 AA.	
XX		
AC	AAU97726;	
XX		
DT	13-AUG-2002 (first entry)	
XX		
DE	Epitope tag #2 used in screening for internalised antibodies.	
XX		
KW	Antibody internalisation; high-throughput screening; cytoplasm;	
KW	ligand internalisation; epitope.	
XX		
OS	Unidentified.	
XX		
PN	WO200233044-A2.	
XX		
PD	25-APR-2002.	
XX		
PF	17-OCT-2001; 2001WO-US32311.	
XX		
PR	18-OCT-2000; 2000US-241279P.	
XX		

PA (REGC) UNIV CALIFORNIA.
 PS
 PI Marks JD, Nielsen UB, Kirpotin DB;
 XX
 XX WPI; 2002-463310/49.
 XX
 XX High-throughput screening for internalising antibodies and identifying
 PT ligands that are internalised into a cell, comprises detecting the
 PT presence of a reporter within the cell that has been contacted with a
 PT ligand -
 XX
 XX Example 1; Page 43; 7lpp; English.
 PS
 XX The invention relates to identifying ligands that are internalised into a
 CC cell comprising detecting the presence of a reporter within the cell that
 CC has been contacted with a ligand, where the presence of the reporter
 CC within the cell indicates that the ligand is internalised into the cell.
 CC Epitope tags are used to non-covalently couple an effector to a ligand.
 CC The method is useful for high-throughput screening for internalising
 CC antibodies and identifying ligands that are internalised into a cell. The
 CC method allows the study of receptor function and the determination of the
 CC temporal-spatial pattern of receptor expression. The cell-specific
 CC receptor ligand, more preferably internalising cell specific receptor
 CC ligand is useful for targeting drugs or markers to the cell surface or
 CC into the cytoplasm (for internalising receptors) e.g. for therapeutic
 CC effect. This sequence represents an epitope tag used in the method of the
 CC invention.
 XX
 XX Sequence 10 AA;
 SQ
 Query Match 83.3%; Score 5; DB 23; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.68;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GHHHH 5
 Db 4 GHHHH 8
 RESULT 21
 ABB76374
 ID ABB76374 standard; Peptide; 10 AA.
 XX
 AC ABB76374;
 XX
 DT 22-AUG-2002 (first entry)
 DE Hexa-histidine tag.
 XX
 XX 6-Hydroxymethyl-7,8-dihydropterin pyrophosphokinase; HPPK;
 KW antibiotic; antagonist; antifolate; alpha/beta protein;
 KW histidine tag.
 XX
 OS Synthetic.
 XX
 PN WO200231511-A2.
 XX
 PD 18-APR-2002.
 XX
 PF 12-OCT-2001; 2001WO-US32047.
 XX
 XX 12-OCT-2000; 2000US-239750P.
 PR
 XX (ICOS-) ICOS CORP.
 PA
 XX Stauton DE;
 PI
 XX WPI; 2002-471361/50.
 DR
 XX Modulating binding interactions between alpha/beta proteins comprising
 PT allosteric regulatory sites, and a binding partner, by contacting with
 PT an allosteric effector molecule that interacts with the regulatory site
 PT -

XX
 PS
 XX Example 19; Page 135; 163pp; English.
 CC
 CC The present sequence is a hexahistidine affinity tag produced as
 CC an N-terminal extension of a recombinant Escherichia coli
 CC 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase (HPPK)
 CC protein produced in an example from the invention. The affinity
 CC tag facilitated purification of the recombinant protein. HPPK is
 CC part of the de novo folate biosynthetic cascade and catalyses the
 CC transfer of pyrophosphate from ATP to 6-hydroxy-7,8-dihydropterin.
 CC It is not expressed in higher eukaryotes, and therefore represents a
 CC novel target for the development of antibiotics with antifolate
 CC activity. Recombinant HPPK protein was used to identify small
 CC molecule inhibitors of HPPK. These inhibitors were capable of
 CC inhibiting the growth of E. coli ToIC (MIC = 0.1-12.5 uM). HPPK is
 CC an example of an alpha/beta protein that comprises a Rossmann fold
 CC structure as an allosteric regulatory site. The invention provides
 CC a method for modulating binding interactions between alpha/beta
 CC proteins and a binding partner using an allosteric effector
 CC molecule.
 XX
 XX Sequence 10 AA;
 SQ
 Query Match 83.3%; Score 5; DB 23; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.68;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GHHHH 5
 Db 2 GHHHH 6
 RESULT 22
 AAU78933
 ID AAU78933 standard; Peptide; 10 AA.
 XX
 AC AAU78933;
 XX
 DT 18-JUN-2002 (first entry)
 DE 9 Histidine peptide.
 XX
 XX Nuclear localisation signal; NLS; protein delivery;
 KW fusion protein; membrane penetrating peptide.
 XX
 OS Synthetic.
 XX
 PN WO200218572-A2.
 XX
 PD 07-MAR-2002.
 XX
 PF 23-AUG-2001; 2001WO-US26421.
 XX
 PR 25-AUG-2000; 2000US-227647P.
 PR 07-FEB-2001; 2001GB-0003110.
 XX
 XX (AVET) AVENTIS PHARM INC.
 PA
 XX Guo Y, Morse CC, Yao Z, Keesler GA;
 PI
 XX WPI; 2002-304256/34.
 DR
 XX New fusion proteins comprising membrane penetrating peptides, useful as
 PT in vivo, ex vivo or in vitro intracellular carriers or delivery devices
 PT for a compound of interest (e.g. peptide, protein, chemical entity,
 PT nucleic acid) -
 XX
 XX Example 2; Page 27; 45pp; English.
 PS
 XX This invention relates to a novel fusion protein, which comprises a
 CC membrane penetrating peptide attached to a compound of interest.
 CC The membrane penetrating peptide of the fusion protein is derived from a
 CC nuclear localisation signal and may be the nuclear localisation signal

CC from human period protein hPER1. The fusion protein is useful for
 CC delivery of a compound of interest into a cell. The fusion protein is
 CC useful as in vivo, ex vivo or in vitro intracellular delivery devices
 CC for a compound of interest (e.g. peptide, protein, chemical entity,
 CC nucleic acid). In particular, the polypeptides are useful as protein
 CC carriers for delivery of compounds to cells. The present sequence
 CC represents the 9 Histidine synthetic peptide used in an assay to analyse
 CC the ability of different peptides to penetrate cellular membranes in
 CC the examples of the invention.

XX Sequence 10 AA;

Query Match 83.3%; Score 5; DB 23; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.68;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHHH 5
 DB 1 GHHHH 5

RESULT 23
 ID AAO16174 standard; Peptide; 10 AA.

XX AAO16174;

DT 06-MAR-2003 (first entry)

XX O6-alkylguanine-DNA alkyltransferase fusion protein-related peptide #2.
 DE Fusion protein; label transfer; AGT-fusion protein; AGT substrate;
 DE O6-alkylguanine-DNA alkyltransferase; AGT; affinity tag transfer.

XX Unidentified.

XX WO200283937-A2.

XX 24-OCT-2002.

XX 05-APR-2002; 2002WO-GB01636.

XX 10-APR-2001; 2001US-282766P.

XX (ECOL-) ECOLE POLYTECHNIQUE FEDERALE LAUSANNE.

XX (KIDD/) KIDDLE S J.

XX Kiddle SJ, Johnsson K, Gendreizig S, Keppeler A;

XX WPI; 2003-103375/09.

XX Transferring label from substrate to fusion protein comprising
 XX O-alkylguanine-DNA alkyltransferase, comprises contacting protein, the
 XX enzyme, and substrate having label, so that enzyme transfers the label
 XX which bonds to the protein -

XX Example A; Page 19; 41pp; English.

XX The invention comprises a method for transferring a label from a
 XX substrate to a fusion protein having O6-alkylguanine-DNA alkyltransferase
 XX (AGT). The method involves contacting a fusion protein containing the
 XX desired protein, AGT and a substance having a label so that AGT transfers
 XX the label and it becomes covalently bonded to the fusion protein. The
 XX method of the invention is useful for transferring a label from a
 XX substrate to a fusion protein using AGT. The method may also be used to
 XX transfer an affinity tag to the AGT-fusion protein. The present amino
 XX acid sequence represents a peptide that was used in the method of the
 XX invention.

XX Sequence 10 AA;

Query Match 83.3%; Score 5; DB 24; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.68;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 HHHS 6
 DB 4 HHHS 8

RESULT 24

ID AAE30843 standard; peptide; 10 AA.

XX AAE30843;

DT 24-FEB-2003 (first entry)

DE HIS tag peptide used to construct hZcytor11/hCRF2-4 heterodimer.

XX Cytokine receptor; Zcytor16; immune function; autoimmune disease;
 KW inflammatory bowel disease; ulcerative colitis; Crohn's disease; asthma;
 KW arthritis; psoriasis; endotoxaemia; septicemia; toxic shock syndrome;
 KW infectious disease; cancer; diabetes; renal dysfunction; pancreatitis;
 KW spontaneous abortion; placental health; gene therapy; nephrotropic.

XX Synthetic.

XX WO200277174-A2.

XX 03-OCT-2002.

XX 22-MAR-2002; 2002WO-US08911.

XX 27-MAR-2001; 2001US-279222P.

XX (ZYMO) ZYMOGENETICS INC.

XX Presnell SR, Xu W, Kindsvogel M, Chen Z, Hughes SD;

XX WPI; 2003-029924/02.

XX New Zcytor16 polypeptide useful for treating or detecting autoimmune or
 PT inflammatory diseases, e.g. inflammatory bowel disease, arthritis,
 PT asthma, diabetes, cancer, psoriasis, endotoxaemia, septicemia, or
 PT infectious diseases -

XX Example 28; Page 264; 268pp; English.

XX The invention relates to cytokine receptor polypeptide, Zcytor16. The
 CC polypeptide is useful in modulating the immune system by binding Zcytor16
 CC ligand and thus preventing the binding of the ligand with endogenous
 CC Zcytor16 receptor. It is useful for studying human inflammation or immune
 CC function, or for treating autoimmune or inflammatory diseases such as
 CC inflammatory bowel disease, ulcerative colitis, Crohn's disease,
 CC arthritis, psoriasis, endotoxaemia, septicemia, toxic shock syndrome or
 CC infectious disease, asthma, cancer, diabetes, renal dysfunctions or
 CC pancreatitis or in assessing therapeutic aspects of IL-11F, chemical
 CC therapeutics, anti-IL-TIF antibodies, anti-Zcytor16 antibodies or
 CC Zcytor16 soluble receptors. The nucleic acid molecule and the antibody
 CC are useful as probes in detecting gene expression and gene structure,
 CC such as in the diagnosis and/or prevention of spontaneous abortions or
 CC in monitoring placental health and function. The invention is useful in
 CC gene therapy. The present sequence is C-terminal HIS tag with spacer
 CC peptide, used in the construction of hZcytor11/hCRF2-4 heterodimer.

XX Sequence 10 AA;

Query Match 83.3%; Score 5; DB 24; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.68;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHS 5

DB 4 GHHS 8

RESULT 25
 AAR95060
 ID AAR95060 standard; Peptide; 4 AA.
 XX AC AAR95060;
 XX 18-AUG-1996 (first entry)
 XX Multidomain protein spacer peptide.
 DE
 XX Nucleic acid transfer system; Gene transfer; gene therapy;
 KW cell targeting; multidomain protein; vector; cancer.
 XX Synthetic.
 OS
 XX WO9613599-A1.
 PN
 XX 09-MAY-1996.
 PD
 XX 31-OCT-1995; 95WO-EP04270.
 PF
 XX 01-NOV-1994; 94EP-0810627.
 PR
 XX (WELS/) WELS W.
 PA
 XX Fominaya J, Wells W;
 PI
 XX WPI; 1996-239505/24.
 DR
 XX Nucleic acid transfer system for gene therapy, e.g. against cancer
 PT - includes toxin translocation domain to target nucleic acid to
 PT specific cell
 XX
 XX Disclosure; Page 6; 106pp; English.
 XX
 CC Spacer peptides (AAR95059-61) are used to link the components of
 CC multidomain proteins (see also AAR95053-58) comprising several
 CC functional domains including a nucleic acid binding domain. The
 CC multidomain proteins are used with an effector nucleic acid, which
 CC includes e.g. a gene and a cognate structure recognizable by the
 CC nucleic acid binding domain of the protein, in a novel nucleic acid
 CC transfer system suitable for gene therapy.
 XX
 SQ Sequence 4 AA;
 Query Match 66.7%; Score 4; DB 17; Length 4;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 HHHH 5
 Db 1 HHHH 4

RESULT 26
 AAW19196
 ID AAW19196 standard; peptide; 4 AA.
 XX AC AAW19196;
 XX 06-AUG-1997 (first entry)
 DT
 XX Isoelectric point marker peptide 55.
 DE
 XX Label: fluorescent dye; isoelectric; point; pi; marker; focusing;
 KW fluorescence; capillary; stable.
 XX Synthetic.
 OS
 XX EP744614-A2.
 PN
 XX 27-NOV-1996.
 PD

XX 29-MAR-1996; 96EP-0105113.
 PF
 XX 19-OCT-1995; 95JP-0271196.
 PR
 XX 31-MAR-1995; 95JP-0076873.
 PR
 XX (MOLE-) LAB MOLECULAR BIOPHOTONICS.
 PA
 XX Kasai K, Matsumoto H, Shimura K, Takamoto H;
 PI
 XX WPI; 1997-001360/01.
 DR
 XX Isoelectric point markers for isoelectric focusing - comprising
 PT fluorescence-labelled oligonucleotide(s)
 PT
 XX Example; Page 11; 29pp; English.
 PS
 XX The present peptide, when labelled with a fluorescent dye, can be
 CC used as an isoelectric point (pi) marker for isoelectric focusing
 CC with fluorescence. The dye is linked to the peptide's
 CC amino-terminal through an amide, thioamide, sulphonamide, urea,
 CC thiourea or urethane bond, and is rhodamine, fluorescein, cyanine,
 CC indocyanine, indocarbocyanine, pyronine, lucifer yellow,
 CC quinarine, squarillium, coumarin, fluoroanthranil maleimide or
 CC anthracene. The marker can be used for capillary isoelectric
 CC focusing, and it is possible to construct peptide sets that cover
 CC a wide pi range and have good storage stability.
 CC The present peptide has a calculated pi value of 6.94.
 XX
 SQ Sequence 4 AA;
 Query Match 66.7%; Score 4; DB 18; Length 4;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 HHHH 5
 Db 1 HHHH 4

RESULT 27
 AAR25918
 ID AAR25918 standard; peptide; 5 AA.
 XX AC AAR25918;
 XX 25-MAR-2003 (updated)
 DT
 XX 21-JAN-1993 (first entry)
 DE
 XX Chelating peptide for immobilising and purifying proteins.
 DE
 XX Inert; spacer; active molecule; CP-IMAC; assay; ligands.
 KW
 XX Synthetic.
 OS
 XX EP497585-A2.
 PN
 XX 05-AUG-1992.
 PD
 XX 30-JAN-1992; 92EP-0300775.
 PF
 XX 30-JAN-1991; 91US-0647901.
 PR
 XX (ANDE/) ANDERSON L D.
 PA (COOK/) COOK J A.
 PA (DAVI/) DAVID G S.
 PA (HOCH/) HOCHSCHWENDER S M.
 PA (KASH/) KASHER M S.
 PA (SMIT/) SMITH M C.
 PA (STEM/) STEMMER W P C.
 XX
 XX Anderson LD, Cook JA, David GS, Hochschwender SM, Kasher MS;
 PI Smith MC, Stemmer WPC;

XX
DR WPI; 1992-261005/32.
XX
XX Immobilisation and purification of molecules - using kinetically
PT inert immobilised metal-CP-protein complex useful in assay
PT system
XX
XX
PS Disclosure; Page 9; 88pp; English.
XX
XX The peptide is a chelating peptide which is used in the complex;
CC biologically active mol.-spacer-chelator-transition metal. The
CC complex may be used in a method which simplifies existing procedures
CC of immobilising and purifying proteins to a single column, by an
CC improvement in the CP-IMAC technique. The kinetically inert complex
CC provides a component of an assay system useful for studying the
CC interaction of any of a variety of ligands with the immobilised
CC chelator peptide-protein. The method takes advantage of the
CC specificity inherent in the CP-IMAC technique, yet opens avenues in
CC the development of assay systems of greater simplicity and
CC effectiveness. Immobilised chelating peptide-derivatised protein
CC complex may be used to isolate and purify antibodies against the
CC protein. Immobilisation of immunoreactive proteins or other bio-
CC logically active molecules enables orientation of these mols. so as
CC to maximise exposure of the antigen or ligand binding site in an
CC affinity chromatography system.
CC See also AAR25916-25, AAR25941-2.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX

SQ Sequence 5 AA;
Query Match 66.7%; Score 4; DB 13; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
Db 1 HHHH 4

RESULT 28
AAR25922
ID AAR25922 standard; peptide; 5 AA.
XX
XX AAR25922;
XX
XX 25-MAR-2003 (updated)
XX 21-JAN-1993 (first entry)
XX
XX Chelating peptide for immobilising and purifying proteins.
XX Inert; spacer; active molecule; CP-IMAC; assay; ligands.
XX Synthetic.
XX EP497585-A2.
XX
XX 05-AUG-1992.
XX
XX 30-JAN-1992; 92EP-0300775.
XX
XX 30-JAN-1991; 91US-0647901.
XX
XX (ANDE/) ANDERSON L D.
XX (COOK/) COOK J A.
XX (DAVI/) DAVID G S.
XX (HOCH/) HOCHSCHWENDER S M.
XX (KASH/) KASHER M S.
XX (SMIT/) SMITH M C.
XX (STEM/) STEMMER W P C.
XX
XX Anderson LD, Cook JA, David GS, Hochschwender SM, Kasher MS;
XX Smith MC, Stemmer WPC;

DR WPI; 1992-261005/32.
XX
XX Immobilisation and purification of molecules - using kinetically
PT inert immobilised metal-CP-protein complex useful in assay
PT system
XX
XX
PS Disclosure; Page 9; 88pp; English.
XX
XX The peptide is a chelating peptide which is used in the complex;
CC biologically active mol.-spacer-chelator-transition metal. The
CC complex may be used in a method which simplifies existing procedures
CC of immobilising and purifying proteins to a single column, by an
CC improvement in the CP-IMAC technique. The kinetically inert complex
CC provides a component of an assay system useful for studying the
CC interaction of any of a variety of ligands with the immobilised
CC chelator peptide-protein. The method takes advantage of the
CC specificity inherent in the CP-IMAC technique, yet opens avenues in
CC the development of assay systems of greater simplicity and
CC effectiveness. Immobilised chelating peptide-derivatised protein
CC complex may be used to isolate and purify antibodies against the
CC protein. Immobilisation of immunoreactive proteins or other bio-
CC logically active molecules enables orientation of these mols. so as
CC to maximise exposure of the antigen or ligand binding site in an
CC affinity chromatography system.
CC See also AAR25916-25, AAR25941-2.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX

SQ Sequence 5 AA;
Query Match 66.7%; Score 4; DB 13; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
Db 1 HHHH 4

RESULT 29
AAR48988
ID AAR48988 standard; Protein; 5 AA.
XX
XX AAR48988;
XX
XX 25-MAR-2003 (updated)
XX 30-AUG-1994 (first entry)
XX
XX Atomic force microscopy complex chelator peptide #6.
XX Chelating peptide; atomic force microscopy; complex; AFM substrate;
XX ME; protein; enzyme; nucleotide; carbohydrate; lipid; receptor;
XX transition metal ion; immobilise; visualisation; substrate; ligand;
XX interaction; enzyme action; antibody-antigen recognition.
XX Synthetic.
XX EP583894-A2.
XX
XX 23-FEB-1994.
XX
XX 26-JUL-1993; 93EP-0305866.
XX
XX 29-JUL-1992; 92US-0922494.
XX
XX (HYBR-) HYBRITECH INC.
XX David GS, Hale JE, Nakamura KK, Richard III C;
XX WPI; 1994-058789/08.
XX
XX New complexes for atomic force microscopy - having a molecular
XX entity bonded via a chelator and a transition metal to an AFM
XX substrate

XX PS Claim 8; Page 25; 33pp; English.

XX CC The sequences given in AAR4893-91 are chelating peptides which were

XX CC used in the atomic force microscopy complexes of the invention. The

XX CC complexes are of the formula:

XX CC ME-(spacer)x-Chelator-(X)-AFM substrate

XX CC AFM substrate = an atomic force microscopy (AFM) substrate;

XX CC ME = a molecular entity, eg. proteins, enzymes, nucleotides,

XX CC carbohydrates or lipids;

XX CC x = 0 or 1;

XX CC M = a transition metal ion, with a proviso that each [ME-(spacer)x-

XX CC chelator-(M)-] unit bound to the AFM substrate may be the same or

XX CC different. The complexes can immobilise a molecular entity on an AFM

XX CC substrate with high specificity in a manner sufficiently stable to

XX CC resist the disrupting forces encountered during interaction with a

XX CC mechanical device such as an AFM tip. The molecular entities can be

XX CC immobilised to allow visualisation of the molecular entities and their

XX CC interaction with their substrate/ligand/receptor. This is particularly

XX CC useful for the study of enzyme action, antibody-antigen recognition,

XX CC etc. for structure based design of systems for in vivo or in vitro use.

XX CC Molecular entities may also be manipulated on a microscopic scale using

XX CC the amplex.

XX CC (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 5 AA;

Query Match 66.7%; Score 4; DB 15; Length 5;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5

DB 1 HHHH 4

RESULT 30

AAR80761

ID AAR80761 standard; peptide; 5 AA.

AC AAR80761;

DT 25-MAR-2003 (updated)

DT 14-MAR-1996 (first entry)

DE Transition metal ion chelating peptide.

DE Chelating agent; affinity chromatography; isolation; diagnosis;

KW ligand.

KW Synthetic.

OS US5439829-A.

PN 08-AUG-1995.

XX 24-JAN-1992; 92US-0826928.

XX 24-JAN-1992; 92US-0826928.

PR 30-JAN-1991; 91US-0647901.

XX (HYBR-) HYBRITECH INC.

PA (ELIL) LILLY & CO ELI.

XX Anderson LD, Cook JA, David GS, Hochschwender SM;

PI Kasher MS, Smith MC, Stemmer WPC;

XX WPI; 1995-283101/37.

XX Binding biologically active cpd., esp. protein, to a support - by

PT coupling to chelating agent, reaction with support carrying

PT transition metal ions then oxidn. or redn. to form kinetically inert

PT oxidn. state, useful e.g. in affinity chromatography

XX PS Claim 8; Column 61; 69pp; English.

XX CC AAR80759-R80768 are transition metal ion chelating peptides capable

XX CC of chelating Co(II), Cr(II) or Ru(III). These peptides are

XX CC covalently bonded to a biologically active compound, esp. a protein,

XX CC and the chelating peptide-protein product is treated with a solid

XX CC support carrying transition metal ions so that the chelating

XX CC peptide-protein conjugate binds the support. The conjugate is then

XX CC treated with either an oxidising or reducing agent to change the

XX CC metal ion oxidn. state to Co(III), Cr(III) or Ru(II), forming a

XX CC kinetically inert oxidn. state so that the biologically active

XX CC compound or protein is immobilised. This new method of

XX CC immobilisation is useful in the preparation of affinity chromatography

XX CC columns used to isolate and purify proteins and antibodies. The

XX CC method can be used in the diagnosis of disease and for the study of

XX CC oncoprotein/anti-oncoprotein interaction (in the design of anticancer

XX CC agents).

XX CC (Updated on 25-MAR-2003 to correct PF field.)

XX SQ Sequence 5 AA;

Query Match 66.7%; Score 4; DB 16; Length 5;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5

DB 1 HHHH 4

RESULT 31

AAR80765

ID AAR80765 standard; peptide; 5 AA.

AC AAR80765;

XX 25-MAR-2003 (updated)

DT 14-MAR-1996 (first entry)

DE Transition metal ion chelating peptide.

DE Chelating agent; affinity chromatography; isolation; diagnosis;

KW ligand.

KW Synthetic.

OS US5439829-A.

PN 08-AUG-1995.

XX 24-JAN-1992; 92US-0826928.

XX 24-JAN-1992; 92US-0826928.

PR 30-JAN-1991; 91US-0647901.

XX (HYBR-) HYBRITECH INC.

PA (ELIL) LILLY & CO ELI.

XX Anderson LD, Cook JA, David GS, Hochschwender SM;

PI Kasher MS, Smith MC, Stemmer WPC;

XX WPI; 1995-283101/37.

XX Binding biologically active cpd., esp. protein, to a support - by

PT coupling to chelating agent, reaction with support carrying

PT transition metal ions then oxidn. or redn. to form kinetically inert

PT oxidn. state, useful e.g. in affinity chromatography

XX Claim 8; Column 63; 69pp; English.

XX CC AAR80759-R80768 are transition metal ion chelating peptides capable

XX CC of chelating Co(II), Cr(II) or Ru(III). These peptides are

CC covalently bonded to a biologically active compound, esp. a protein,
 CC and the chelating peptide-protein product is treated with a solid
 CC support carrying transition metal ions so that the chelating
 CC peptide-protein conjugate binds the support. The conjugate is then
 CC treated with either an oxidising or reducing agent to change the
 CC metal ion oxidn. state to Co(III), Cr(III) or Ru(II), forming a
 CC kinetically inert oxidn. state so that the biologically active
 CC compound or protein is immobilised. This new method of
 CC immobilisation is useful in the preparation of affinity chromatography
 CC columns used to isolate and purify proteins and antibodies. The
 CC method can be used in the diagnosis of disease and for the study of
 CC oncoprotein/anti-oncoprotein interaction (in the design of anticancer
 CC agents).
 CC (Updated on 25-MAR-2003 to correct PF field.)

SQ Sequence 5 AA;
 Query Match 66.7%; Score 4; DB 16; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DY 2 HHHH 5
 DB 1 HHHH 4

RESULT 32
 AAW19194
 ID AAW19194 standard; peptide; 5 AA.
 AC AAW19194;
 XX
 DT 06-AUG-1997 (first entry)
 DE Isolelectric point marker peptide 53.
 KW Label; fluorescent dye; isoelectric; point; pI; marker; focusing;
 QW fluorescence; capillary; stable.
 CS Synthetic.
 XS
 SN EP744614-A2.
 XD 27-NOV-1996.
 IX 29-MAR-1996; 96EP-0105113.
 X 19-OCT-1995; 95JP-0271196.
 R 31-MAR-1995; 95JP-0076873.
 CR (MOLE-) LAB MOLECULAR BIOPHOTONICS.
 CA Kasai K, Matsumoto H, Shimura K, Takamoto H;
 CR WPI; 1997-001360/01.
 TX Isoelectric point markers for isoelectric focusing - comprising
 T fluorescence-labelled oligo:nucleotide(s)
 X Example; Page 11; 29pp; English.

CC The present peptide, when labelled with a fluorescent dye, can be
 CC used as an isoelectric point (pI) marker for isoelectric focusing
 CC with fluorescence. The dye is linked to the peptide's
 CC amino-terminal through an amide, thioamide, sulphonamide, urea,
 CC thiourea or urethane bond, and is rhodamine, fluorescein, cyanine,
 CC indocyanine, indocarbocyanine, pyronine, lucifer yellow,
 CC quinarine, squarillium, coumarin, fluoranthranil maleimide or
 C anthracene. The marker can be used for capillary isoelectric
 C focusing, and it is possible to construct peptide sets that cover
 C a wide pI range and have good storage stability.
 C The present peptide has a calculated pI value of 7.06.

SQ Sequence 5 AA;
 Query Match 66.7%; Score 4; DB 18; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 HHHH 5
 DB 1 HHHH 4
 RESULT 33
 AAW81153
 ID AAW81153 standard; peptide; 5 AA.
 AC AAW81153;
 XX
 DT 25-MAR-2003 (updated)
 DT 21-JUN-1999 (first entry)
 XX
 DE Chelating peptide (c).
 XX Penicillin binding protein; PBP2A-27R; methicillin resistance;
 KW antibiotic; assay; purification; chelating peptide.
 XX Synthetic.
 XX EP875578-A2.
 XX 04-NOV-1998.
 XX 18-MAR-1992; 98EP-0200530.
 XX 19-MAR-1991; 91US-0672704.
 PR 18-MAR-1992; 92EP-0302298.
 XX (ELIL) LILLY & CO ELI.
 XX Blaszcak LC, Skatrud PL, Smith MC, Wu CE;
 PI WPI; 1998-559443/48.
 DR New Staphylococcus aureus soluble penicillin-binding proteins and
 PT their derivatives - useful for screening for compounds effective
 PT against methicillin resistant organisms
 XX Disclosure; Page 25; 97pp; English.
 PS The invention provides 8 chelating peptides (see AAW81151-58) of the
 CC formula (His)x-(A)y-(His)z, where A is an amino acid, x = 1-10,
 CC y = 0-4, and z = 1-10, as well as monomers, dimers and trimers of
 CC these peptides in which each monomer unit is the same or different.
 CC The chelating peptides are joined, optionally via a Pro residue, to
 CC the N-terminal end of novel soluble penicillin binding proteins
 CC (PBPs) of the invention (see AAW81159-62), preferably derived from
 CC PBP2A-27R (see AAW81149), the protein responsible for the methicillin
 CC resistance of Staphylococcus aureus strain 27R. The new PBP2A-27R
 CC proteins are useful for assaying for agents useful as antibiotics
 CC against methicillin resistant Staphylococcus strains. Soluble
 CC forms of PBP2A-27R facilitate crystallisation as they lack their
 CC transmembrane association region, and so are useful for x-ray
 CC crystallography studies of the protein, assisting in the design of
 CC antibiotic compounds. The chelating peptide operably linked to the
 CC PBP2A-27R proteins facilitates purification, while the optional Pro
 CC residue placed between the chelating peptide and the PBPs provides
 CC a dipeptidase digestion site that facilitates removal of the
 CC chelating peptide.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 CC (Updated on 25-MAR-2003 to correct PR field.)
 XX
 SQ Sequence 5 AA;

Query Match 66.7%; Score 4; DB 19; Length 5;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
Db 1 HHHH 4

RESULT 34

AAW81157
ID AAW81157 standard; peptide; 5 AA.

XX AC AAW81157;
XX 25-MAR-2003 (updated)
DT 21-JUN-1999 (first entry)

XX DE Chelating peptide (g).

XX KW Penicillin binding protein; PBP2A-27R; methicillin resistance;
XX KW antibiotic; assay; purification; chelating peptide.
XX OS Synthetic.
XX PN EP875578-A2.
XX PD C4-NOV-1998.
XX PF 18-MAR-1992; 98EP-0200530.
XX PR 19-MAR-1991; 91US-0672704.
XX PR 18-MAR-1992; 92EP-0302298.
XX PA (ELIL) LILLY & CO ELI.
XX PI Blaszcak LC, Skatrud PL, Smith MC, Wu CE;
XX DR WPI; 1998-559443/48.

XX New Staphylococcus aureus soluble penicillin-binding proteins and
PT their derivatives - useful for screening for compounds effective
PT against methicillin resistant organisms

XX Claim 3; Page 47; 97pp; English.

XX The invention provides 8 chelating peptides (see AAW81151-58) of the
CC formula (His)_x(A)_y(His)_z, where A is an amino acid, x = 1-10,
CC y = 0-4, and z = 1-10, as well as monomers, dimers and trimers of
CC these peptides in which each monomer unit is the same or different.
CC The chelating peptides are joined, optionally via a Pro residue, to
CC the N-terminal end of novel soluble penicillin binding proteins
CC (PBPs) of the invention (see AAW81159-62), preferably derived from
CC PBP2A-27R (see AAW81149), the protein responsible for the methicillin
CC resistance of Staphylococcus aureus strain 27R. The new PBP2A-27R
CC proteins are useful for assaying for agents useful as antibiotics
CC against methicillin resistant Staphylococcus strains. Soluble
CC forms of PBP2A-27R facilitate crystallization as they lack their
CC transmembrane association region, and so are useful for x-ray
CC crystallography studies of the protein, assisting in the design of
CC antibiotic compounds. The chelating peptide operably linked to the
CC PBP2A-27R proteins facilitates purification, while the optional Pro
CC residue placed between the chelating peptide and the PBPs provides
CC a dipeptidase digestion site that facilitates removal of the
CC chelating peptide.
CC (Updated on 25-MAR-2003 to correct PF field.)
CC (Updated on 25-MAR-2003 to correct PR field.)

XX Sequence 5 AA;

Query Match 66.7%; Score 4; DB 19; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
Db 1 HHHH 4

RESULT 35

AAE11984
ID AAE11984 standard; peptide; 5 AA.

XX AC AAE11984;
XX 18-DEC-2001 (first entry)

XX DE Tag #1, to facilitate conjugation of variant with non-polypeptide moiety.
XX KW Interferon-gamma; IFNG; interferon-beta; interferon alpha; FSH;
KW follicle stimulating hormone; granulocyte colony stimulating factor;
KW G-CSF; interleukin 10; IL-10; medicament; pulmonary administration;
KW circulatory administration.
XX OS Synthetic.
XX PN WO200168141-A2.
XX PD 20-SEP-2001.
XX PF 16-MAR-2001; 2001WO-DK00182.
XX PR 17-MAR-2000; 2000DK-0000447.
XX PR 20-MAR-2000; 2000US-190844P.
XX PR 25-AUG-2000; 2000WO-DK00471.
XX PR 13-NOV-2000; 2000WO-DK00631.
XX PA (MAXY-) MAXYGEN APS.
XX PI Hansen CK;
XX DR WPI; 2001-616274/71.

XX Dispersions of particles for administration to the lung and circulatory
PT system by inhalation, e.g. from a nebulizer -
XX Disclosure; Page 35; 139pp; English.

XX The present invention relates to novel dispersions of a conjugate
CC comprising at least one non-polypeptide moiety covalently attached
CC to a polypeptide variant, wherein the amino acid sequence of the
CC polypeptide variant differs from that of the corresponding wild-type
CC human polypeptide in that at least 1 residue comprising an attachment
CC group for the non-polypeptide group has been introduced and/or removed.
CC The wild-type human polypeptide is selected from the group consisting
CC of interferon-gamma (IFNG), interferon-beta, interferon alpha, follicle
CC stimulating hormone (FSH), interleukin 10 (L-10) and granulocyte colony
CC stimulating factor (G-CSF). The dispersion is used for the preparation
CC of a medicament for pulmonary and/or circulatory administration. The
CC present peptide sequence is a tag which facilitates fast and easy
CC purification of the conjugate. This tag serves as a convenient tool
CC for achieving conjugation between the tagged variant and the non-
CC polypeptide moiety. The tag may function as a spacer molecule ensuring
CC an improved accessibility to the immobilised variant to be conjugated.

XX Sequence 5 AA;

Query Match 66.7%; Score 4; DB 22; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
Db 1 HHHH 4

RESULT 36

XX 19-OCT-1995; 95JP-0271196.
 PR 31-MAR-1995; 95JP-0076873.
 XX (MOLE-) LAB MOLECULAR BIOPHOTONICS.
 XX Kasai K, Matsumoto H, Shimura K, Takamoto H;
 XX WPI; 1997-001360/01.
 DR Isolelectric point markers for isoelectric focusing - comprising
 PT fluorescence-labelled oligo:nucleotide(s)
 XX Example; Page 11; 29pp; English.
 XX The present peptide, when labelled with a fluorescent dye, can be
 CC used as an isoelectric point (pI) marker for isoelectric focusing
 CC with fluorescence. The dye is linked to the peptide's
 CC amino-terminal through an amide, thioamide, sulphonamide, urea,
 CC thiocrea or urethane bond, and is rhodamine, fluorescein, cyanine,
 CC indocyanine, indocarbocyanine, pyronine, lucifer yellow,
 CC quinacrine, squarillium, coumarin, fluoroanthranil maleimide or
 CC anthracene. The marker can be used for capillary isoelectric
 CC focusing, and it is possible to construct peptide sets that cover
 CC a wide pI range and have good storage stability.
 CC The present peptide has a calculated pI value of 6.94.
 XX Sequence 6 AA;

Query Match 66.7%; Score 4; DB 18; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
 DB 2 HHHH 5

RESULT 39
 AAW63024
 ID AAW63024 standard; Peptide; 6 AA.
 AC AAW63024;
 XX 09-NOV-1998 (first entry)
 DT Hexa-histidine tag.
 DE Dectin-1; dendritic cell; c-type lectin; mouse; immunity; adjuvant;
 KW allergy; autoimmune disease; gene therapy; vaccine; diagnosis;
 KW drug screening; hexa-histidine tag.
 XX Synthetic.
 OS WO9828332-A2.
 PN 02-JUL-1998.
 PD 22-DEC-1997; 97WO-US23761.
 PF 20-DEC-1996; 96US-0772440.
 PR (TEXA) UNIV TEXAS SYSTEM.
 XX Ariizumi K, Takashima A;
 PI WPI; 1998-377594/32.
 DR Nucleic acid encoding dendritic cell specific peptide(s) dectin-1
 PT and -2 - useful, e.g. to regulate immune response, as vaccine
 PT adjuvants, for diagnosis and drug screening
 XX Claim 92; Page 157; 200pp; English.

XX This hexa-histidine tag peptide is preferred for use in fusions
 CC with murine dectin-1 (see AAW69236) or dectin-2 (see AAW69237)
 CC polypeptides of the invention. The dectins are novel murine
 CC dendritic cell (DC) specific members of the C-type lectin family
 CC that are essential for DC-mediated T cell activation. The
 CC His-tagged polypeptides have been expressed in E. coli host cells
 CC using recombinant vector pQE30, and used for production of
 CC antibodies. The invention provides: dectin-1 and dectin-2
 CC polypeptides and DNA sequences, expression vectors, host cells,
 CC probes and primers, antibodies, compounds that modulate
 CC dectin-mediated activation of T cells, transgenic animals, and
 CC dectin ligands.

XX Sequence 6 AA;

Query Match 66.7%; Score 4; DB 19; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
 DB 1 HHHH 4

RESULT 40
 AAW69961
 ID AAW69961 standard; peptide; 6 AA.

AC AAW69961;

DT 08-OCT-1998 (first entry)

DE Poly-His tag for constructing soluble huRANKL.

KW RANK; necrosis factor-kappa B; NF-kB; receptor activator; human;
 KW immune response; inflammatory response; toxic shock; sepsis; TNF;
 KW RANKL; RANK ligand; tumour necrosis factor.

OS Synthetic.

PN WO9828426-A2.

PD 02-JUL-1998.

PF 22-DEC-1997; 97WO-US23775.

PR 14-OCT-1997; 97US-0064671.

PR 23-DEC-1996; 96US-0059978.

PR 07-MAR-1997; 97US-0813509.

XX (IMMV) IMMUNEX CORP.

XX Anderson DM, Galibert LJ, Maraskovsky E;

XX WPI; 1998-377657/32.

XX New isolated ligand for receptor activator of NF-kappa B - used to
 PT develop products for augmenting an immune response for inhibiting an
 PT inflammatory response and for protection of cells
 XX Example 15; Page 67; 80pp; English.

XX This represents a poly-histidine tag used in the construction of a
 CC soluble huRANKL (RANK ligand). RANK (receptor activator of necrosis
 CC factor-kappaB (NF-kB)) is a member of the tumour necrosis factor (TNF)
 CC family. A soluble RANK may be used for inhibiting activation of
 CC NF-kB, by contacting a cell expressing membrane-associated RANK with a
 CC soluble RANK which binds to RANK ligand (RANKL). RANKL polypeptides can
 CC activate RANK and can be used to induce maturation of dendritic cells and
 CC enhance their allo-stimulatory capacity, thereby augmenting an immune
 CC response. The soluble RANK polypeptide composition may also be used for
 CC regulating an immune or inflammatory response. Inhibition of NF-kB by

CC RANK antagonists may be useful in ameliorating negative effects of an
 CC inflammatory response that result from triggering of RANK, e.g. in
 CC treating toxic shock or sepsis, graft-versus-host reactions, or acute
 CC inflammatory reactions. They can also be used in adjunct therapy for
 CC disease characterised by neoplastic cells that express RANK. RANKL
 CC polypeptides can also be used to identify inhibitors of RANK and thus
 CC inhibitors of an inflammatory response, and also for protecting RANK-
 CC expressing cells from the negative effects of chemotherapy or the
 CC presence of high levels of TNF-alpha. The products can also be used for
 CC detection and drug screening.

XX Sequence 6 AA;

Query Match 66.7%; Score 4; DB 19; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
 DB 1 HHHH 4

RESULT 41
 ID AAW68297 standard; peptide; 6 AA.

XX AAW68297;

XX 08-OCT-1998 (first entry)

XX Poly-His tag for constructing soluble huRANKL.

XX RANK; necrosis factor-kappa B; NF-kB; receptor activator; human;
 KW immune response; inflammatory response; toxic shock; sepsis; TNF;
 KW RANKL; RANK ligand; tumour necrosis factor.

XX Synthetic.

XX WO9828424-A2.

XX 02-JUL-1998.

XX 22-DEC-1997; 97WO-US23866.

XX 14-OCT-1997; 97US-0064671.

XX 23-DEC-1996; 96US-0059978.

XX 07-MAR-1997; 97US-0813509.

XX (IMMV) IMMUNEX CORP.

XX Anderson DM, Galibert LJ, Maraskovsky E;

XX WPI; 1998-377655/32.

XX New isolated receptor activator of necrosis factor-kappa B - useful
 for, e.g. developing products for regulating an immune or
 inflammatory response, treating toxic shock or sepsis

XX Example 15; Page 67; 80pp; English.

XX This represents a poly-histidine tag used in the construction of a
 CC soluble huRANKL (RANK ligand). RANK (receptor activator of necrosis
 CC factor-kappaB (NF-kB)) is a member of the tumour necrosis factor (TNF)
 CC family. Host cells transfected or transfected with an expression vector
 CC comprising the RANK encoding nucleic acid can be used to produce
 CC recombinant RANK protein. The soluble RANK may be used for inhibiting
 CC activation of NF-kB, by contacting a cell expressing membrane-associated
 CC RANK with a soluble RANK which binds to RANK ligand (RANKL). The soluble
 CC RANK polypeptide composition may also be used for regulating an immune
 CC or inflammatory response. Inhibition of NF-kB by RANK antagonists may be
 CC useful in ameliorating negative effects of an inflammatory response that
 CC result from triggering of RANK, e.g. in treating toxic shock or sepsis.
 CC graft-versus-host reactions, or acute inflammatory reactions. They can

CC also be used in adjunct therapy for disease characterised by neoplastic
 CC cells that express RANK. The products can also be used for detection and
 CC drug screening.

XX Sequence 6 AA;

Query Match 66.7%; Score 4; DB 19; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
 DB 1 HHHH 4

RESULT 42
 ID AAW44011 standard; peptide; 6 AA.

XX AAW44011;

XX 05-JUN-1998 (first entry)

XX Poly-histidine peptide used in an epitope tagged prion protein construct.

XX Prion; epitope; FLAG; Strep; poly-histidine; haemagglutinin;
 KW recombinant; transgenic animal; scrapie; Creutzfeldt-Jakob disease;
 KW CJD; bovine spongiform encephalopathy; BSE.

XX Synthetic.

XX WO9746572-A1.

XX 11-DEC-1997.

XX 29-MAY-1997; 97WO-US09289.

XX 06-JUN-1996; 96US-0660626.

XX (REGC) UNIV CALIFORNIA.

XX Cohen FE, Prusiner SB, Scott MR, Telling GC;

XX WPI; 1998-042112/04.

XX Nucleic acid construct encoding biologically active protein and
 PT epitope - especially epitope-tagged prion protein

XX Claim 4; Page 48; 62pp; English.

XX This is the sequence of an artificial poly-histidine peptide epitope. It
 CC is used in a recombinant nucleic acid construct encoding an epitope-
 CC tagged prion protein (PrP). The construct comprises a first nucleic acid
 CC sequence encoding an amino acid sequence of a biologically active protein
 CC fragment and a second nucleic acid sequence encoding a heterologous
 CC epitope domain. The heterologous epitope domain is a peptide selected
 CC from a peptide group of FLAG, Strep, poly-histidine, human c-myc peptide
 CC recognised by monoclonal antibody 9E10 and haemagglutinin peptide
 CC synthetic or chimeric PrP molecule. The protein is a natural,
 CC three-dimensional conformation. The protein has two different
 CC positions relative to the protein such that the epitope domain is spatially
 CC exposed in a first conformation relative to a second conformation. The
 CC nucleic acid construct may be used for the production of transgenic
 CC animals or cells that are useful in a method for distinguishing between
 CC different conformational shapes of a protein. These methods are
 CC particularly useful in studying diseases caused by prion proteins, e.g.
 CC Creutzfeldt-Jakob disease (CJD), scrapie and bovine spongiform
 CC encephalopathy (BSE).

XX Sequence 6 AA;

Query Match 66.7%; Score 4; DB 19; Length 6;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
Db 1 HHHH 4

RESULT 43
AA49221
ID AAY49221 standard; peptide; 6 AA.
AC AAY49221;
XX 07-FEB-2000 (first entry)
DT
XX Nickel ion binding domain containing 6 His residues.
DE
XX Monoclonal antibody; MAB; 1A7; GD2; immune response; melanoma;
KW neuroblastoma; glioma; soft tissue carcinoma; small cell carcinoma;
KW tumor-associated antigen.
XX Synthetic.
OS
XX US5977316-A.
PN
XX 02-NOV-1999.
PD
XX 16-JAN-1996; 96US-0591196.
PF
XX 17-JAN-1995; 95US-0372676.
PR
XX (KENT) UNIV KENTUCKY.
PA
XX Room KA, Chatterjee SK, Chatterjee M;
PI WPI; 1999-619711/53.
DR
XX Monoclonal antibody 1A7 which elicits an anti-GD2 immunological
PT response, useful for the development of products for the detection and
PT treatment of cancers .
PT
XX Example 7; Column 58; 74pp; English.
PS
XX The invention provides a monoclonal antibody (MAB) designated 1A7, which
CC elicits an anti-GD2 (tumor-associated antigen) immunological response in
CC humans. MAB 1A7 has defined light and heavy chain variable region
CC sequences. The MAB 1A7 and polypeptides can be used for eliciting an
CC anti-GD2 immune response. The polypeptides can also be used for detecting
CC or purifying anti-GD2 antibody. The products can be used for treating GD2
CC -associated diseases, e.g. melanoma, neuroblastoma, glioma, soft tissue
CC carcinoma, and small cell carcinoma. They can be used for palliating the
CC disease or for reducing the risk of recurrence.
SQ Sequence 6 AA;

Query Match 66.7%; Score 4; DB 20; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
Db 1 HHHH 4

RESULT 44
AAY50711
ID AAY50711 standard; peptide; 6 AA.
AC AAY50711;
XX
XX 20-MAR-2003 (updated)
DT 08-FEB-2000 (first entry)

XX Single chain antibody construct polyhistidine tag.
DE
XX Immunoglobulin; light chain; VL region; heavy chain; VH region;
KW single-chain; antigen binding; variable domain; anticancer; treatment;
KW antiviral; antibacterial; antimalarial; antiinflammatory; diagnosis;
KW tumor vaccine; autoimmune disease; inflammation; blood disorder;
KW nervous system; infection.
XX
OS Unidentified.
XX DE19827239-A1.
PN
XX 23-DEC-1999.
PD
XX 18-JUN-1998; 98DE-1027239.
PF
XX 18-JUN-1998; 98DE-1027239.
PR
XX (HYRI) HOECHST WARION ROUSSEL DEUT GMBH.
PA Kontermann R, Sedlacek H, Mueller R;
XX WPI; 1999-591691/06.
PI
XX Single chain molecule binding antigen, its preparation and medicine
DR containing this molecule - consists of binding some antigen with
XX different variable domain of light and heavy chain of immunoglobulin.
PT
XX Example 1; Page 17; 26pp; German.
PS
XX This invention describes a novel single-chain molecule (I) that binds
CC multiple antigens and comprises two variable domains of heavy
CC immunoglobulin chains (VH) and two variable domains of light chains (VL).
CC The domains are provided as two VH-VL constructs which are attached via
CC a peptide (P). Any VH and VL may be replaced by their functional
CC fragments. The products of the invention have anticancer, antiviral,
CC antibacterial, antimalarial, and antiinflammatory activity. (I) are used
CC to treat, prevent or diagnose tumors (e.g. as tumor vaccines), autoimmune
CC diseases and inflammation (e.g. transplant rejection and arthritis),
CC blood disorders (e.g. of the coagulation and/or circulatory systems, such
CC as anemia, leucopenia, thrombocytopenia and hypertension), nervous system
CC disorders and/or infections (by viruses or bacteria, or malaria,
CC including, when (i) include a fusogenic peptide, use for gene transfer.
CC This sequence represents a polyhistidine tag which is used in the
CC construction of the single chain construct described in the invention.
CC NOTE: This specification is a treat as basic for CZ-9901215 in Derwent
CC week 9951.
CC (Updated on 20-MAR-2003 to correct DR field.)
SQ Sequence 6 AA;

Query Match 66.7%; Score 4; DB 20; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
Db 1 HHHH 4

RESULT 45
AAY33592
ID AAY33592 standard; Protein; 6 AA.
AC AAY33592;
XX
XX 20-DEC-1999 (first entry)
DT
XX VH-VL domain linker peptide #4.
DE
XX Antigen binding; single chain; variable domain; VH domain; light chain;
KW heavy immunoglobulin chain; VL domain; anticancer; antiviral; tumor;

PS Disclosure; Page 7; 116pp; English.

XX This peptide comprises an epitope tag that can be used in
 CC methods of the invention. The invention discloses a powerful new
 CC system for engineering antibody affinity and specificity, by
 CC constructing a microbial analogue of the mammalian system's B cell
 CC repertoire. Antibodies are displayed on the surface of yeast cells
 CC by genetic fusion with yeast cell wall proteins, especially
 CC agglutinin proteins. After mutation, variants are selected on the
 CC basis of improved binding characteristics with fluorescently
 CC labeled targets. The selection method also identifies proteins
 CC with enhanced phenotypic characteristics, proteins that are
 CC displayed at higher levels, proteins that are secreted at higher
 CC efficiency and proteins of improved stability.

XX Sequence 6 AA;

Query Match 66.7%; Score 4; DB 20; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
 DB 1 HHHH 4

RESULT 48

AY17447
 ID AAY17447 standard; peptide; 6 AA.

XX
 AC AAY17447;

XX 28-JUL-1999 (first entry)

DE Major outer membrane protein amino terminal peptide #5.

XX Major outer membrane protein; MOMP; Chlamydia trachomatis; antibody;
 KW antigen.

OS Synthetic.

XX JP1123078-A.

XX 11-MAY-1999.

XX 22-OCT-1997; 97JP-0289938.

XX 22-OCT-1997; 97JP-0289938.

XX (ELED) DENKI KAGAKU KOGYO KK.

XX WPI; 1999-340514/29.

XX New major outer membrane protein of Chlamydia Trachomatis - useful
 PT for determination of Chlamydia Trachomatis antibody

XX Claim 2; Page 7; 11pp; Japanese.

XX The present invention describes a major outer membrane protein (MOMP) of
 CC Chlamydia trachomatis in which six amino acid residues are peptide-bound
 CC to the amino terminal and which has immunological activity. Also
 CC described are: (1) a method for the preparation of MOMP of Chlamydia
 CC trachomatis in which the gene of the above MOMP of Chlamydia trachomatis
 CC is inserted to an expression vector used to transform E. coli, the
 CC transformant is cultured under conditions where the MOMP of Chlamydia
 CC trachomatis is expressed and the expression product is collected from
 CC the transformant; (2) a method for the determination of the antibody of
 CC MOMP of Chlamydia trachomatis using the above MOMP of Chlamydia
 CC trachomatis as the antigen; and (3) a reagent for the determination of
 CC the antibody of MOMP of Chlamydia trachomatis using the above major MOMP
 CC of Chlamydia trachomatis as the antigen. The method can provide an
 CC antigen used for the determination of Chlamydia trachomatis antibody.

SQ Sequence 6 AA;

Query Match 66.7%; Score 4; DB 20; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
 DB 1 HHHH 4

RESULT 49

AY09544
 ID AAY09544 standard; peptide; 6 AA.

XX
 AC AAY09544;

XX 21-JUL-1999 (first entry)

XX Human TAK1 6xHis peptide.

XX Human; TAK1; TAK1; screening; inhibition; TGF-beta;
 KW transforming growth factor beta.

XX Homo sapiens.

XX WO9921010-A1.

XX 29-APR-1999.

XX 22-OCT-1998; 98WO-JP04796.

XX 22-OCT-1997; 97JP-0290188.

XX (CHUS) CHUGAI SEIYAKU KK.

XX Ontomo T, Ono K, Tsuchiya M;

XX WPI; 1999-312645/26.

XX Screening for TGF- beta inhibitory substances, which are useful as
 PT drugs for treatment of diseases relating to its disorder

XX Example 1; Page 58; 195pp; Japanese.

XX A method has been developed for screening for substances which inhibit
 CC the binding of TAK1 polypeptide to TAK1 polypeptide. The method
 CC comprises: (a) contacting the polypeptide in the presence of a sample;
 CC and (b) detecting the amount of bound polypeptide, in which the sample
 CC can be pre-mixed with TAK1 or TAK1 polypeptide first. The transforming
 CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
 CC indications e.g. as TGF-beta signal transduction inhibitors or
 CC activators, or extracellular matrix protein production enhancers
 CC inhibitors or activators, or cell proliferation prevention inhibitors or
 CC activators, or monocyte migration inhibitors or activators, or
 CC physiological activity induction inhibitors or activators, or
 CC immunosuppression inhibitors or activators, or amyloid beta protein
 CC precipitation inhibitors or activators, and such substances can also be
 CC inhibitors of the TAK1 polypeptide function, particularly kinase
 CC activity. The present sequence represents a peptide from an example
 CC of the present invention.

XX Sequence 6 AA;

Query Match 66.7%; Score 4; DB 20; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
 DB 1 HHHH 4

RESULT 50

AAW88019
ID AAW88019 standard; peptide; 6 AA.
XX
AC AAW88019;
XX
09-APR-1999 (first entry)
DT
XX Peptide used in the hair composition of the invention.
DE
XX Hair; styling; wave.
KW
XX
XX
XX
OS Synthetic.
XX
JP11012138-A.
PN
XX
19-JAN-1999.
PD
XX
18-JUN-1997; 97JP-0161657.
PF
XX
18-JUN-1997; 97JP-0161657.
PR
XX
(LLOYD) LION CORP.
PA
XX
WPI; 1999-148442/13.
DR
XX
XX New composition for hair - comprises oligopeptide having more than 2
PT cysteine residues and reducing agent
XX
XX Disclosure; Page 4; 17pp; Japanese.

XX Peptides AAW88006-32 are used in the hair composition of the invention.
CC The specification describes a hair composition that comprises at
CC least one oligopeptide having more than two cysteine residues and more
CC than three amino acid residues of the same kind (except cysteine) and a
CC reducing agent. The composition is useful for styling hair to take
CC various shapes and waves safely and effectively.
XX
XX

Sequence 6 AA;

Query Match 66.7%; Score 4; DB 20; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 2 HHHH 5

b 2 HHHH 5

RESULT 51

AAW84203
ID AAW84203 standard; Peptide; 6 AA.
XX
AC AAW84203;
XX
25-MAR-1999 (first entry)
JT
XX
XX Peptide comprising a purification tag.

XX Proteinase site; bone morphogenetic fusion protein; bone binding site;
CW bone morphogenetic protein; transforming growth factor beta;
CW active fragment; wound healing; bone growth; purification tag.
XX
XX Synthetic.
XX

WO9855137-A1.

X 10-DEC-1998.

X F 02-JUN-1998; 98WO-US11189.
X X
R 03-JUN-1997; 97US-0868452.

XX
PA (HALL/) HALL F L.
PA (HANE/) HAN B.
PA (NINN/) NIMNI M E.
PA (SHOR/) SHORS E C.
PA (WULL/) WJ L.
XX
PI Hall FL, Han B, Nimni ME, Shors EC, Wu L;
XX
XX WPI; 1999-059875/05.
DR N-PSDB; AAV99371.
XX
PT New bone morphogenetic fusion proteins - comprising a purification
PT tag and a bone morphogenetic active fragment, used for enhancing
PT wound healing or bone growth
XX
PS Claim 5; Page 38; 64pp; English.
XX
CC The present sequence represents a peptide comprising a purification tag
CC that was used in the creation of the bone morphogenetic fusion proteins
CC of the invention. The bone morphogenetic fusion protein may contain some
CC or all of the following elements: a purification tag, a proteinase site,
CC an ECM/bone binding site, a second proteinase site, and a bone
CC morphogenetic protein active fragment. The fusion proteins of the
CC invention also includes proteins that have transforming growth factor
CC beta active fragments instead of bone morphogenetic protein active
CC fragments. The bone morphogenetic fusion proteins can be used for
CC enhancing wound healing or bone growth.
XX
XX

Sequence 6 AA;

Query Match 66.7%; Score 4; DB 20; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5

Db 1 HHHH 4

RESULT 52

ABB08438
ID ABB08438 standard; Peptide; 6 AA.
XX
AC ABB08438;
XX
07-MAY-2002 (first entry)
DT
XX
DE Histidine tag relative to the farnesyltransferase of the invention.

KW Farnesyltransferase; enzyme.

OS Unidentified.

XX

XX

XX

XX

XX

XX 01-APR-1997; 97KR-0012067.

XX 01-APR-1997; 97KR-0012067.

XX (GLDS) LG CHEM LTD.

XX Moon GD, Kim MJ, Chung HH;

XX WPI; 2000-020309/02.

XX N-PSDB; ABA98900.

XX Farnesyltransferase having histidine tag and process for preparing the
PT same.
XX
XX Disclosure; Page 23; 23pp; Korean.

XX The invention relates to a farnesyltransferase with a histidine tag,
 CC and methods for preparing it. The current sequence represents a histidine
 CC tag relative to the farnesyltransferase of the invention.
 XX
 SQ Sequence 6 AA;
 Query Match 66.7%; Score 4; DB 21; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 2 HHHH 5
 DB 1 HHHH 4
 ||||
 RESULT 53
 AAB10788
 ID AAB10788 standard; Peptide; 6 AA.
 XX
 AC AAB10788;
 XX
 DT 05-FEB-2001 (first entry)
 XX
 DE B. subtilis lumazine synthase protein fragment #15.
 XX
 XX Lumazine synthase; capsid; cytostatic; antiviral; antibacterial; vaccine;
 KW gene therapy; immunotherapy; biosensor; diagnosis.
 XX
 OS Bacillus subtilis.
 XX
 PN WO2000053229-A2.
 XX
 XX 14-SEP-2000.
 PD
 XX
 XX 03-MAR-2000; 2000WO-EP01899.
 XX
 XX 08-MAR-1999; 99DE-1010102.
 XX
 XX (FISC/) FISCHER M.
 PA (BACH/) BACHER A.
 XX
 XX Fischer M, Bacher A;
 PI WPI; 2000-572230/53.
 XX
 XX Protein conjugate based on lumazine synthase as carrier, useful e.g.
 PT for vaccination and immunotherapy, contains many functional molecules
 PT attached to outside of the carrier -
 XX
 XX Claim 43b; Page 100; 180pp; German.
 PS
 XX This invention describes a novel protein conjugate (I) comprising at
 CC least 1 functional region (FR) at any position in the sequence of a
 CC carrier protein (II) to form a capid three-dimensional structure of a
 CC lumazine synthase (LS) type, such that the outer periphery is covalently
 CC linked to many FR. The invention also describes (1) a hetero-oligomeric
 CC protein conjugate (Ia) comprising either a mixture of at least 2
 CC different (I) or of at least 1 (I) and at least 1 (II) that lacks FR,
 CC with the components optionally covalently linked by chemical treatment;
 CC (2) a method for preparation of (I) and (Ia); (3) a vector for
 CC preparation of (I); (4) a DNA (III) that encodes (I); (5) LS from
 CC Bacillus subtilis with Cys 93 and/or Cys139 replaced by Ser; (6) DNA
 CC (IIa) encoding LS from Aquifex aeolicus which is codon-optimized for
 CC expression in a recombinant Escherichia coli strain; (7) a chimeric
 CC protein (CP) comprising amino acids (aa) 1-60 from LS of B. subtilis and
 CC aa 61-154 from LS of A. aeolicus for use as (II); and (8) pharmaceutical
 CC compositions and vaccines containing (I) and (Ia). The products of the
 CC invention have cytostatic, antiviral and antibacterial activity and can
 CC be used in a vaccine or for gene therapy. (I) are used as pharmaceuticals
 CC (e.g. for immunotherapy of human immune deficiency virus infection or
 CC tumors) and in vaccines, including multivalent vaccines, against
 CC bacterial or viral infections, to produce diagnostic or therapeutic

CC antibodies, for selective detection, purification and characterization of
 CC antibodies, and for preparation of protein libraries. (I) may also be
 CC used as biosensors and for diagnosis, e.g. of tumors. Genes that encode
 CC (I) are useful in DNA vaccines and for preparation of plant-based oral
 CC vaccines. (I) can contain many FR (same or different) at the surface of a
 CC spherical particle (LS comprises 60 subunits that assemble into an
 CC icosahedron). The large number of FR may increase sensitivity in
 CC immunoassays and the efficiency of immunotherapy agents.
 XX
 SQ Sequence 6 AA;
 Query Match 66.7%; Score 4; DB 21; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 2 HHHH 5
 DB 1 HHHH 4
 ||||
 RESULT 54
 AAY96951
 ID AAY96951 standard; peptide; 6 AA.
 XX
 AC AAY96951;
 XX
 DT 31-OCT-2000 (first entry)
 XX
 DE Polyhistidine tag for use in rubredoxin fusion protein.
 XX
 KW PRUBEX3; polyhistidine; tag; rubredoxin; antigen; vaccine; carrier;
 KW immunogenic; flag tag; fusion.
 XX
 OS Synthetic.
 XX
 PN WO2000039310-A1.
 XX
 PD 06-JUL-2000.
 XX
 XX 29-DEC-1999; 99WO-US31176.
 XX
 XX 29-DEC-1998; 98US-0114034.
 PR
 XX (UYGE-) UNIV GEORGIA RES FOUND INC.
 XX
 XX Przybyla A, Menon N;
 PI WPI; 2000-452403/39.
 DR
 XX Nucleic acids encoding fusion peptides comprising rubredoxin, useful as
 PT vaccines
 PT
 XX Claim 16; Page 61; 67pp; English.
 PS
 XX Fusion proteins comprising rubredoxin and fused antigens and/or haptens
 CC may be used as vaccines to initiate immune responses. In this case, the
 CC rubredoxin acts as a carrier. The fusion protein is capable of binding
 CC iron (Fe-2+) when properly folded, giving it a red color that makes it
 CC easy to identify following or during purification. The C-terminal fused
 CC protein may be insoluble or known to form inclusion bodies in a host
 CC cell. The rubredoxin serves as the carrier molecule to yield immunogenic
 CC fusion products. As rubredoxin is itself only negligibly antigenic, there
 CC is no need to include a cleavage site in the fusion product to allow
 CC cleavage of the N-terminal and C-terminal constituents.
 XX
 SQ Sequence 6 AA;
 Query Match 66.7%; Score 4; DB 21; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 2 HHHH 5
 DB 1 HHHH 4
 ||||

PD	20-JAN-2000.	
XX	13-JUL-1999; 99WO-US15870.	
XX	13-JUL-1998; 98US-0092671.	
XX	(CELL-) CELLOMICS INC.	
XX	Rubin RA, Giuliano KA, Gough A, Dunlay T;	
XX	WPI; 2000-171170/15.	
XX	Automated screening method for identifying compounds which induce cell	
XX	surface receptor internalization, useful for drug discovery -	
XX	Example 6; Page 67; 148pp; English.	
XX	The invention relates to a method for identifying compounds which	
XX	inhibit internalization of cell surface receptors. Provided are an array	
XX	of locations, each containing cells with a cell surface receptor	
XX	protein, that are treated with a test compound. The protein is	
XX	luminescently labeled or contacted with a luminescently labeled cell	
XX	before or after test compound treatment. Any luminescence produced is	
XX	converted into digital data and automatically analysed to determine if	
XX	the test compound induced the protein internalization. The novel method	
XX	is used to screen for compounds which modulate cell surface receptor	
XX	protein internalization, this can be used in drug discovery to test	
XX	compound efficacy in living biological systems. The assay method is	
XX	automated and compact. It has high throughput and uses smaller volumes of	
XX	reagents and test compounds. Sequences AAY77704-718 represent examples of	
XX	peptide epitope tags used in the course of the invention.	
XX	Sequence 6 AA;	
XX	Query Match 66.7%; Score 4; DB 21; Length 6;	
XX	Best Local Similarity 100.0%; Pred. No. 9.3e+05;	
XX	Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
XX	2 HHHH 5	
XX	1 HHHH 4	
XX	RESULT 57	
XX	AAY78346	
XX	ID AAY78346 standard; peptide; 6 AA.	
XX	AAAY78346;	
XX	05-MAY-2000 (first entry)	
XX	His-6 peptide SEQ ID NO:2.	
XX	Immunoglobulin; protein capture; detection; antibody; proteomic;	
XX	drug screening; diagnosis.	
XX	Synthetic.	
XX	WO200004389-A2.	
XX	27-JAN-2000.	
XX	14-JUL-1999; 99WO-US15968.	
XX	14-JUL-1998; 98US-0115455.	
XX	(ZYOM-) ZYOMYX INC.	
XX	Wagner P, Nock S, Ault-Riche D, Itin C;	
XX	WPI; 2000-161175/14.	
XX	New arrays for assaying proteins, used for analysis of cell expression	

PT products, evaluating disease conditions, proteomics, drug screening,
 PT diagnostics and measurement of gene activity -
 XX
 PS Example 6; Page 68; 90pp; English.

XX The present invention describes an array of protein-capture agents (PCA),
 CC comprising a substrate, at least one organic thinfilm covering some or
 CC all of the surface of the substrate, and patches arranged in discrete,
 CC known regions on the portions of the substrate surface covered by organic
 CC thinfilm. Each patch comprises PCAs, capable of binding a particular
 CC expression product, or a fragment of a cell population, immobilised on
 CC the organic thinfilm. The array comprises different PCAs, capable of
 CC binding different expression products, or fragments, of the cell
 CC population. The arrays can be used for assaying for expression products
 CC or fragments of a cell or population of cells. They can be used for
 CC evaluating a disease condition in a tissue of an organism. They can be
 CC used in proteomics, drug screening, diagnostics, and the measurement of
 CC gene activity at the protein level in cells. The present sequence
 CC represents a peptide used in an example from the present invention.

XX Sequence 6 AA;

Query Match 66.7%; Score 4; DB 21; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 HHHH 5
 ||||
 Db 1 HHHH 4

RESULT 58
 AAY78351
 ID AAY78351 standard; peptide; 6 AA.

XX

AC AAY78351;

XX 05-MAY-2000 (first entry)

DT His-6 peptide SEQ ID NO:2.

DE Caspase; protein array; screening; biomolecular activity; proteomic;

XX drug development; biosensor; diagnosis.

XX Synthetic.

XX WO200004382-A1.

XX 27-JAN-2000.

XX 14-JUL-1999; 99WO-US15971.

XX 14-JUL-1998; 98US-0115455.

XX (ZYOM-) ZYOMYX INC.

XX Wagner P, Ault-Riche D, Nock S, Itin C;

XX WPI; 2000-171289/15.

XX New arrays for analyzing components of a fluid sample, useful for drug
 PT development, functional proteomics, clinical diagnostics and biosensors
 PS Example 9; Page 58; 81pp; English.

XX The present invention describes an array (I) of proteins comprising a
 CC substrate, at least one organic thinfilm on some or all of the substrate
 CC surface, and patches arranged in discrete, known regions on portions of
 CC the substrate surface covered by organic thinfilm. Each of the patches
 CC comprises a protein immobilised on the underlying organic thinfilm.
 CC The arrays can be used for screening proteins for their ability to
 CC interact with a component of a sample. They can also be used for
 CC assaying for protein-protein binding interactions or analytes. They can

CC be used for drug development, proteomics, clinical diagnostics and
 CC biosensors. The present sequence represents a peptide used in the
 CC exemplification of the present invention.

SQ Sequence 6 AA;

Query Match 66.7%; Score 4; DB 21; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 HHHH 5
 ||||
 Db 1 HHHH 4

RESULT 59

ABR00924

ID ABR00924 standard; Peptide; 6 AA.

XX ABR00924;

XX 03-APR-2003 (first entry)

DT Hexahistidine peptide tag SEQ ID NO 32.

DE Human; Gene 216; chromosome 20p13-p12; antiasthmatic; anorectic;
 KW antiinflammatory; gastrointestinal; gene therapy; vaccine; asthma;
 KW obesity; inflammatory bowel disease.

XX Synthetic.

XX WO200178894-A2.

XX 25-OCT-2001.

XX 13-APR-2001; 2001WO-US12245.

XX 13-APR-2000; 2000US-0548797.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Keith T;

XX WPI; 2001-639428/73.

XX Isolated genes (Gene 216) from human chromosome 20p13-p12 and the
 PT proteins they encode, useful for the prevention, diagnosis and
 PT treatment of asthma, obesity and inflammatory bowel disease -

XX Disclosure; Page 45; 520pp; English.

XX The invention relates to isolated genes (Gene 216) from human chromosome
 CC 20p13-p12 and the proteins they encode. The nucleic acids and proteins
 CC may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate Gene 216 expression. For example, the
 CC nucleic acids (or vectors) and proteins may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of gene 216 by expressing
 CC inactive proteins or to supplement the patients own production of Gene
 CC 216 proteins. Additionally, the nucleic acids may be used to produce the
 CC secreted Gene 216 protein, by inserting the nucleic acids into a host
 CC cell and culturing the cell to express the protein. The nucleic acids
 CC and complementary sequences may also be used as DNA probes in diagnostic
 CC assays to detect and quantitate the presence of similar nucleic acid
 CC sequences in samples and therefore which patients may be in need of
 CC restorative therapy. The Gene 216 protein may also be used as antigens in
 CC the production of antibodies against Gene 216 and in assays to identify
 CC modulators of Gene 216 expression and activity. The anti-Gene 216
 CC antibodies and antagonists may also be used to down regulate expression
 CC and activity. The anti-Gene 216 antibodies may also be used as diagnostic
 CC agents for detecting the presence of Gene 216 proteins in samples (e.g.
 CC by enzyme linked immunosorbant assay or ELISA). Disorders that may be
 CC prevented, diagnosed and/or treated by the above methods include, for

CC example asthma, obesity and inflammatory bowel disease. The present
 CC sequence is that of a peptide tag used in analysis of the Gene 216
 CC encoded protein.

XX SQ Sequence 6 AA;

Query Match 66.7%; Score 4; DB 22; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
 DB 1 HHHH 4

RESULT 60
 ABB76820
 ID ABB76820 standard; peptide; 6 AA.

AC ABB76820;

XX 16-JUL-2002 (first entry)

DT His tag.

DE Angiotensin-Y2; His tag.

XX Unidentified.

OS KR2000059883-A.

PN 16-OCT-2000.

PD 09-MAR-1999; 99KR-0007779.

PF 09-MAR-1999; 99KR-0007779.

PR (KOHG/) KOH G Y.

PA (KIMI/) KIM I J.

XX Kwak HJ, Ahn JE, Koh GN, Kim SY;

PI WPI; 2001-254061/26.

XX Novel angiotensin-Y2 gene and recombinant protein - NoAbstract

PT Example 1; Page 4; 10pp; Korean.

FS The present invention relates to a novel angiotensin-Y2 gene and

XX recombinant protein. The present sequence is a His tag used in an example

CC from the invention.

XX SQ Sequence 6 AA;

Query Match 66.7%; Score 4; DB 22; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5

DB 1 HHHH 4

RESULT 61

AAM52173

ID AAM52173 standard; peptide; 6 AA.

XX AAM52173;

AC AAM52173;

XX 07-FEB-2002 (first entry)

DE Peptide tag 1.

KW Factor VII; FVII; Factor VIIa; FVIIa; haemostatic; thrombolytic;
 KW cardiant; hepatotropic; cerebroprotective; haemophilia; liver disease;
 KW myocardial infarction; thrombotic stroke; deep-vein thrombosis.

XX Synthetic.

XX WO200158935-A2.

XX 16-AUG-2001.

XX 12-FEB-2001; 2001WO-DK00094.

XX 11-FEB-2000; 2000DK-0000218.

XX 18-OCT-2000; 2000DK-0001558.

XX (MAXY-) MAXYGEN APS.

XX Andersen KV, Pedersen AH, Bornaes C;

XX WPI; 2001-581807/65.

XX New conjugate, useful for treating Factor VIIa related diseases or
 P disorders such as haemophilia, liver disease, myocardial infarction and
 P deep-vein thrombosis, comprises non-polypeptide group covalently
 PT attached to polypeptide group -

PS Disclosure; Page 42; 89pp; English.

XX The invention relates to novel Factor VII (FVII) or Factor VIIa (FVIIa)
 CC polypeptide conjugates, comprising at least one non-polypeptide group
 CC covalently attached to a polypeptide, where the amino acid sequence of
 CC polypeptide differs from that of the wildtype FVIIa (AAM52171) in that at
 CC least one amino acid residue containing an attachment group for the
 CC non-polypeptide group has been introduced or removed. The FVIIa
 CC conjugates have haemostatic, thrombolytic, cardiant, hepatotropic and
 CC cerebroprotective activity and are useful for treating FVIIa/TF-related
 CC diseases or disorders such as haemophilia, liver disease, myocardial
 CC infarction, thrombotic stroke and deep-vein thrombosis. The conjugates
 CC have increased functional in vivo half life and/or increased plasma half
 CC life, increased bioavailability and or reduced sensitivity to proteolytic
 CC degradation. Consequently medical treatment using the conjugates has a
 CC number of advantages over currently available such as longer duration
 CC between injections. The present sequence is that of a peptide tag, useful
 CC to the invention.

XX SQ Sequence 6 AA;

Query Match 66.7%; Score 4; DB 22; Length 6;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5

DB 1 HHHH 4

RESULT 62

AAE13079

ID AAE13079 standard; peptide; 6 AA.

XX AAE13079;

AC AAE13079;

XX 28-JAN-2002 (first entry)

DE Epitope tag #7 used in yeast cell surface display of proteins.

XX Phenotypic property; yeast; cell wall protein; epitope tag.

XX Unidentified.

XX US6300065-B1.

XX 09-OCT-2001.

XX 26-AUG-1998; 98US-0140084.
 PF 31-MAY-1996; 96US-018741P.
 PR 30-MAY-1997; 97US-0866398.
 PR 20-JAN-1998; 98US-0009388.
 XX (UNII) UNIV ILLINOIS FOUND.
 XX Kieke MC, Wittrup KD, Boder ET, Kranz DM, Shusta E;
 PI WPI; 2001-656236/75.
 DR Selecting proteins, e.g. antibodies, with enhanced phenotypic
 XX properties relative to those of a wild-type comprises transforming
 PT yeast cells with a vector expressing a protein to be tested fused to a
 PT yeast cell wall protein -
 XX Disclosure; Column 5; 64pp; English.
 PS The present invention relates to a method for selecting proteins with
 XX enhanced phenotypic properties relative to those of a wild-type,
 CC comprises transforming yeast cells with a vector expressing a protein
 CC to be tested fused to a yeast cell wall protein. The method is
 CC particularly useful for selecting antibodies for improved affinity and
 CC specificity. The present sequence is an epitope tag which is used in
 CC yeast cell surface display of proteins.
 XX Sequence 6 AA;
 SQ Query Match 66.7%; Score 4; DB 22; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 HHHH 5
 DB 1 HHHH 4
 RESULT 63
 AAMS1720
 ID AAMS1720 standard; peptide; 6 AA.
 XX AAMS1720;
 AC 29-JAN-2002 (first entry)
 XX FSH alpha or beta tagging peptide 1.
 DE Human; FSH alpha; FSH beta; follicle stimulating hormone; glycosylation;
 XX antifertility.
 KW Synthetic.
 XX WO200158493-A1.
 PN 16-AUG-2001.
 XX 09-FEB-2001; 2001WO-DK00090.
 XX 11-FEB-2000; 2000DK-0000220.
 PR 14-JUL-2000; 2000DK-0001092.
 XX (MAXY-) MAXYGEN APS.
 PA Schambye HT, Andersen KV, Van Den Hazel B, Christiansen J;
 XX Jeppesen CB;
 PI WPI; 2001-607186/69.
 DR New polypeptide conjugate with follicle stimulating hormone (FSH)
 XX PT activity, used to treat infertility, comprises polypeptide having
 PT modified FSH alpha and beta subunits with attachment group for

PT non-polypeptide moiety -
 XX Disclosure; Page 37; 88pp; English.
 PS The invention relates to new polypeptides and polypeptide conjugates
 CC exhibiting follicle stimulating hormone (FSH) activity. FSH is a dimeric
 CC hormone comprising of an alpha (AAMS1709) and beta (AAMS1711) subunit.
 CC The invention relates to a heterodimeric FSH conjugate comprising, a
 CC dimeric polypeptide having FSH alpha (AAMS1733-AAMS1800,
 CC AAMS1442-AAMS1449) and beta (AAMS2001-AAMS2104) subunits, where at least
 CC one subunit differs from corresponding wild-type subunit in that an
 CC amino acid residue containing an attachment group for a non-polypeptide
 CC molecule has been introduced or removed, especially where at least one of
 CC the FSH-alpha and FSH-beta subunits comprises at least one introduced
 CC N- or O-glycosylation site at its N-terminal and the glycosylation site
 CC being glycosylated. The polypeptides have antifertility activity. The
 CC polypeptides have increased functional in vivo half life and/or serum
 CC half life as compared to human FSH, replenishing insufficient endogenous
 CC FSH production in a patient. The present sequence is that of a peptide
 CC tag for protein purification.
 XX Sequence 6 AA;
 SQ Query Match 66.7%; Score 4; DB 22; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 HHHH 5
 DB 1 HHHH 4
 RESULT 64
 AAE12713
 ID AAE12713 standard; peptide; 6 AA.
 XX AAE12713;
 AC 04-JAN-2002 (first entry)
 XX Hexahis tag used to construct human bivPH1-IL-2 fusion protein.
 DE Tumour-associated antigen mucin-1; MUC-1; adenocarcinoma; cancer;
 XX breast; ovary; lung; bladder; cytostatic; therapy; PH1 antibody.
 KW Synthetic.
 XX WO200175110-A2.
 PN 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US10589.
 PF 30-MAR-2000; 2000US-0538913.
 PR (DYAX-) DYAX CORP.
 XX Hoogenboom HRJM, Henderikx MPG;
 PI WPI; 2001-626437/72.
 DR N-PSDB; AAD20743.
 XX Novel isolated tumor-associated antigen mucin-1-specific binding member
 PT for diagnosing and treating cancer, comprises mucin-1 binding domain or
 PT its portion for binding to an epitope of the protein core of mucin-1 -
 XX Example 1; Page 36; 126pp; English.
 PS The invention relates to an isolated tumour-associated antigen mucin-1
 XX (MUC-1)-specific binding member comprising an antigen binding domain
 CC region having an antibody variable light (VL) or heavy (VH) region,
 CC or a complementarity determining region (CDR) of VL or VH. MUC1-specific
 CC binding member is useful for diagnosing cancer, preferably adenocarcinoma

CC The binding of MUC1-specific binding member to MUC1 is detected by a
 CC detection method selected from enzyme-linked immunosorbent assay,
 CC magnetic resonance imaging, scintillation counting, and X-ray film.
 CC MUC1-specific binding member is useful for treating cancer, preferably
 CC adenocarcinoma, in an individual, where the cancer is present in tissue
 CC of the breast, ovary, lung, or bladder of the individual. MUC1-specific
 CC binding member is useful for diagnosing and imaging MUC1-expressing
 CC cancer cells and tissues, for purifying or isolating non-glycosylated,
 CC underglycosylated or cancer-associated forms of MUC1 or MUC1 epitope-
 CC containing molecules, and for therapeutically or prophylactically
 CC treating cancer. The present sequence is a polyhistidine (hexahis) tag
 CC used for constructing human bivpH1-IL-2 fusion protein which is an
 CC immunocytokine MUC1-specific binding member.

XX Sequence 6 AA;

Query Match 66.7%; Score 4; DB 22; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 HHHH 5
 b 1 HHHH 4

RESULT 65

AG65669
 D AAG65669 standard; peptide; 6 AA.

XX C AAG65669;

XX T 07-JAN-2002 (first entry)

E His tag used for recombinant expression of FGF-like polypeptide.

X Fibroblast growth factor; FGF; FGF-like polypeptide; mitogenic;
 W fat deposition; vulvar; antiulcer; dermatological; anorectic;
 W antidiabetic; antiinflammatory; cytostatic; hepatic; virucide;
 W neuroprotectant; pulmonary; gene therapy; vaccine; human.

S Synthetic.

X N WO200172957-A2.

X D 04-OCT-2001.

X F 02-APR-2001; 2001WO-IB00664.

X R 31-MAR-2000; 2000US-0540118.

X X (ITOH/) ITOH N.

X A Itoh N;

X R WPI; 2001-611623/70.

T New human nucleic acid encoding fibroblast growth factor-like peptide,
 T useful for treatment and diagnosis of e.g. wounds and inflammatory
 T bowel disease -

S Example 3; Page 117; 172pp; English.

X The invention provides human nucleic acids encoding fibroblast growth
 C factor (FGF)-like peptide. The FGF-like polypeptides can be expressed by
 C standard recombinant methodology and are mitogenic for a wide range of
 C cells, inducing differentiation and proliferation, and inhibiting
 C deposition of fat. The FGF-like polypeptides, polynucleotides and
 C specific antibodies and modulators are useful for treating a very wide
 C range of diseases and conditions, e.g. wounds, ulcers, skin aging,
 C obesity, diabetes, alopecia, inflammatory bowel disease, emphysema, viral
 C hepatitis, multiple sclerosis, respiratory distress syndrome, tumors of
 C the eye, etc., also for maintaining organs before transplant and
 C supporting culture of primary cells and tissues. Sequences AAG65668-69

CC represent amino acid sequences of protein tags used for recombinant
 CC expression of FGF-like polypeptide.

XX Sequence 6 AA;

Query Match 66.7%; Score 4; DB 22; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
 Db 1 HHHH 4

RESULT 66

AAG79069
 ID AAG79069 standard; peptide; 6 AA.

XX AC AAG79069;

XX DT 10-DEC-2001 (first entry)

XX DE Peptide which inhibits vascular endothelial growth factor (VEGF).

KW Vascular endothelial growth factor; VEGF; VEGF inhibitor; cancer;
 KW angiogenesis-related disease; diabetic retinopathy; rheumatoid arthritis.
 XX Synthetic.

XX FN WO200166127-A1.

XX PD 13-SEP-2001.

XX PF 21-DEC-1999; 95WO-KR00796.

XX PR 21-DEC-1999; 99WO-KR00796.

XX PA (GREC) KOREA GREEN CROSS CORP.

XX PA (POST-) POSTECH FOUND.

XX PI Chae CB, Bae DG, Yoon WH;

XX DR WPI; 2001-602600/68.

XX PT New arginine-rich peptides, useful as vascular endothelial growth
 PT factor inhibitors for treating cancers and other angiogenesis-related
 PT diseases such as rheumatoid arthritis and diabetic retinopathy -

XX PS Disclosure; Page 10; 65pp; English.

XX CC The present sequence represents a peptide from a hexa-peptide
 CC library, which was tested for its ability to inhibit the activity
 CC of vascular endothelial growth factor (VEGF). Peptides of the invention
 CC which inhibit VEGF comprise six amino acid residues with arginine at
 CC the first, the fourth and the sixth positions from the amino end, one
 CC selected from arginine, lysine, and histidine at the second position,
 CC and one selected from arginine and lysine at the third and the fifth
 CC positions. The peptides inhibit the binding of VEGF to its receptors.
 CC The peptides inhibit the growth of host normal cells (vascular
 CC endothelial cells), but not cancer cells themselves, and thus overcome
 CC the problems of conventional therapies for cancer, which are due to the
 CC versatility and resistance of cancer cells. The VEGF-inhibiting peptides
 CC are used for treating cancer and angiogenesis-related diseases. They are
 CC also used for inhibiting the growth and metastasis of cancer cells.
 CC Angiogenesis-related diseases include diabetic retinopathy and rheumatoid
 CC arthritis.

XX Sequence 6 AA;

Query Match 66.7%; Score 4; DB 22; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHH 4
DB 1 GHHH 4

RESULT 67
AAE08742 standard; peptide: 6 AA.
AAE08742;
15-NOV-2001 (first entry)
Poly-His tag, to generate soluble RANKL.

Receptor activator of nuclear factor kappaB ligand; RANKL; NF; tumour necrosis factor; TNF; TNF receptor associated factor; TRAF; immune response; inflammatory response; graft-versus-host reaction; toxic shock; sepsis; acute inflammatory reaction; bone resorption; anti-apoptotic signal; therapy; immunosuppressant; anti-inflammatory.

Synthetic.
US6271349-B1.
07-AUG-2001.
17-DEC-1998; 98US-0215649.
23-DEC-1996; 96US-0059978.
07-MAR-1997; 97US-0077181.
14-OCT-1997; 97US-0064671.
23-DEC-1996; 96US-0772330.
07-MAR-1997; 97US-0813509.
22-DEC-1997; 97US-0996139.

(IMMV) IMMUNEX CORP.
Dougall WC, Galibert L;
WPI; 2001-52013/57.

New receptor activator of NF-kappaB (RANK) polypeptides, useful for regulating immune response, in screening for RANK inhibitors, or as an adjunct therapy for disease characterized by neoplastic cells that express RANK -

Example 15; Column 81; 47pp; English.

The patent discloses novel receptor activator of nuclear factor (NF) - kappaB (RANK) proteins and their corresponding DNAs. RANK is a member of the tumour necrosis factor (TNF) receptor superfamily and associates with TNF receptor associated factor (TRAF) 2 and 3 which are important in the regulation of immune and inflammatory response. The receptors are useful for regulating immune response and in screening for inhibitors of these receptors. The cytoplasmic domain of RANK is used in developing assays for inhibitors of signal transduction, e.g. for screening the molecules that inhibit interaction of RANK with TRAF1, TRAF2, TRAF3, TRAF5 and particularly TRAF6. NF-kappaB inhibition by RANK antagonists are useful in ameliorating the negative effects of an inflammatory response that result from triggering of RANK, e.g. in treating toxic shock or sepsis, graft-versus-host reactions, acute inflammatory reactions and the effects of bone resorption. RANK acts as an anti-apoptotic signal and rescue the cells that express RANK from apoptosis. Soluble forms of the receptor are used in vivo or in vitro based screening tests for agonists or antagonists of RANK activity, as antagonists of RANK-mediated NF-kappa B activation, or to inhibit transduction of a signal via RANK. RANK compositions are used in the development of both agonistic and antagonistic antibodies, or as an adjunct therapy for disease characterised by neoplastic cells that express RANK. Compounds that interfere with RANK/TRAF6 interactions are useful for modulating the formation of osteoclasts from osteoclast precursors and for modulating osteoclast function and activities. They

CC are used as inhibitors of diseases associated with excess bone resorption and as immunosuppressants or anti-inflammatory agents. The RANK DNAs are useful for the expression of recombinant proteins, as probes for analysis of the presence or distribution of RANK transcripts, while the proteins are useful in preparing kits for the detection of soluble RANK, or CC monitor RANK-related activity. The present sequence is a poly-His tag CC which is used in the exemplification of the invention to generate a CC soluble, tagged, poly-His version of human and murine RANKL.

XX
SQ Sequence 6 AA;
Query Match 66.7%; Score 4; DB 22; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 1 HHHH 4

RESULT 68
AAG78043 standard; peptide: 6 AA.
AAG78043;
20-NOV-2001 (first entry)
6-His epitope tag.

abnormal base-pairing; polymorphism; nucleic acid repair enzyme; mutation; infection; cancer; immune system disorder; metabolic disorder; muscle disorder; bone disorder; nervous system disorder; high throughput screening; epitope tag.

Synthetic.
WO200162968-A2.
30-AUG-2001.
05-JAN-2001; 2001WO-US00452.
25-FEB-2000; 2000US-0514016.
(GEAT) GEN ATOMICS.
Yuan C;
WPI; 2001-536643/59.

Detecting abnormal base-pairing, mutation in nucleic acid, or polymorphism in gene locus, comprises contacting nucleic acid with abnormal base-pairing and mutant nucleic acid repair enzyme, and detecting their binding -

Disclosure; Page 180; 294pp; English.

The invention relates to detecting abnormal base-pairing in a nucleic acid duplex, mutation in a nucleic acid or polymorphism in a gene locus, comprising contacting a nucleic acid duplex having an abnormal base-pairing with a mutant nucleic acid repair enzyme or its complex and detecting the binding between the nucleic acid duplex and mutant enzyme, such that the presence of abnormal base-pairing, mutation or polymorphism is detected. The method is useful for prognosis or diagnosis of the presence or severity of a disease, disorder or infection by a pathological agent associated with the mutation, including cancer, immune system disorders, metabolic disorders, muscle and bone disorders, nervous system disorders, signal disorders and transporter disease or disorder. The method is rapid and accurate and is amenable to high throughput formats. The method requires neither specific probes nor gel electrophoresis and is amenable to automation for simultaneous detection of a large number of nucleic acid mutations. The present

CC sequence is that of an epitope tag, useful to the invention.

CC
SQ Sequence 6 AA;
Query Match 66.7%; Score 4; DB 22; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 1 HHHH 4

RESULT 69
AAB85699
ID AAB85699 standard; peptide; 6 AA.

CC AAB85699;
KW 29-OCT-2001 (first entry)
DE Amino acid sequence of six-histidine epitope tag.
KW Multivalent protein; immune response; Plasmodium vivax; parasite;
KW protozoacide; vaccine; malaria; recombinant; ViVac1; ViVac2.

CC Synthetic.

CC WO200155181-A2.
CC 02-AUG-2001.

CC 29-JAN-2001; 2001WO-US02937.

CC 31-JAN-2000; 2000US-0179213.

CC (USSH) US DEPT HEALTH & HUMAN SERVICES.

CC Lal AA, Xiao L, Zhou Z;

CC WPI; 2001-514557/56.

CC New recombinant multivalent protein comprising antigenic determinants
CC derived from more than one stage in a life cycle of Plasmodium vivax,
CC useful as a vaccine for treating, preventing and reducing malarial
CC infection -

CC Example 1; Page 25; 59pp; English.

CC The invention relates to recombinant multivalent proteins (I) that
CC stimulate an immune response to Plasmodium vivax. (I) comprises antigenic
CC determinants, fragments or conservative substitutions, derived from more
CC than one stage in a life cycle of a Plasmodium vivax parasite. (I) is
CC useful as a vaccine for stimulating an immune response, specifically a
CC protective immune response that confers increased resistance to infection
CC by Plasmodium parasites, such as P. vivax. (I) is especially useful in
CC the treatment, prevention and reduction of malarial infection, as
CC research or diagnostic reagents for the detection of Plasmodium species
CC in a biological sample, and for conferring immunity against multiple
CC stages of the malarial parasite. The antibodies produced are useful for
CC the detection or measurement of antigenic epitopes derived from one or
CC more stages in a life cycle of a parasite, particularly P. vivax. The
CC vaccine comprising the recombinant proteins, is cost-effective, health-
CC promoting intervention for controlling, preventing or treating the
CC incidence of malaria. The present sequence represents the amino acid
CC sequence of a six-histidine epitope tag, a component of the multivalent
CC and multistage proteins ViVac1p and ViVac2p.

CC Sequence 6 AA;

Query Match 66.7%; Score 4; DB 22; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 1 HHHH 4

RESULT 70
AAE08006

ID AAE08006 standard; peptide; 6 AA.

CC AAE08006;

CC 01-NOV-2001 (first entry)

DE His tag useful as antibody recognition sequences.

CC Neuropeptide Y; NPY receptor; G-protein-coupled transmembrane protein;
CC transmembrane; TM domain; therapy; obesity; blood pressure; epilepsy;
CC Huntington's disorder; Parkinson's disorder; eating disorder; seizure;
CC locomotor; anxiety disorder; limbic seizure; tranquilliser; his tag.

CC Unidentified.

CC WO200155103-A2.

CC 02-AUG-2001.

CC 29-JAN-2001; 2001WO-US02904.

CC 28-JAN-2000; 2000US-0178652.

CC (NEUR-) NEUROGEN CORP.

CC Bennett M, Brodbeck R, Krause J;

CC WPI; 2001-514543/56.

CC New chimeric receptor proteins comprising a single polypeptide chain of
CC amino acids, useful as targets for drug actions, and as basis for drug
CC discovery and development -

CC Disclosure; Page 54; 72pp; English.

CC The present invention relates to chimeric neuropeptide Y (NPY) receptors.
CC The NPY receptors are G-protein-coupled transmembrane proteins with
CC seven membrane spanning transmembrane (TM) domains. The compounds that
CC modulate the activity of a NPY receptor is useful in the preparation of
CC a medicament for treating conditions including obesity, high/low blood
CC pressure, epilepsy, Huntington's and Parkinson's disorder and eating,
CC seizure, locomotor and anxiety disorders. They can also be used as
CC targets for drug actions, and as basis for drug discovery and
CC development. The NPY5 receptor may have an anti-epileptic activity in
CC the control of limbic seizures. The present sequence is a
CC hexa-histidine (His-6x) tag which is used as antibody recognition
CC sequences in the exemplification of the invention.

CC Sequence 6 AA;

Query Match 66.7%; Score 4; DB 22; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 1 HHHH 4

RESULT 71
AAG62605

ID AAG62605 standard; peptide; 6 AA.

CC AAG62605;

DT 06-SEP-2001 (first entry)
 XX Metal capturing protein related peptide #2.
 DE
 XX Metal capturing protein; metal capture; secretory signal;
 KW waste treatment.
 XX
 OS Synthetic.
 XX
 PN WO200138517-A1.
 XX
 XX 31-MAY-2001.
 XX
 XX 26-OCT-2000; 2000WO-JP07518.
 PF
 XX 19-NOV-1999; 95JP-0330226.
 PR
 XX (TOYT) TOYOTA JIDOSHA KK.
 PA
 XX Tanaka A, Ueda M;
 PI
 XX WPI; 2001-355927/37.
 DR
 DR N-PSDB; AAH45709.
 XX
 PT Fused gene with DNA expressing polypeptide capable of capturing metal,
 PT for recombinant vectors and transformants applicable in purifying
 PT environment and recovering metal efficiently, including waste treatment
 PT
 XX
 PS Claim 4; Page 32; 45pp; Japanese.
 XX
 CC The present invention relates to a fused gene containing DNAs encoding a
 CC secretory signal peptide, a protein capable of capturing a metal and a
 CC protein localised on the cell surface. The gene can be used to express
 CC the metal capturing protein, which can then be used in purifying and
 CC recovering metal for example in waste treatment. The present sequence is
 CC a peptide described in the exemplification of the invention.
 CC
 XX
 SQ Sequence 6 AA;
 Query Match 66.7%; Score 4; DB 22; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 HHHH 5
 Db ||||
 1 HHHH 4
 RESULT 72
 AAU02090
 ID AAU02090 standard; Peptide; 6 AA.
 XX
 AC AAU02090;
 XX
 DT 07-SEP-2001 (first entry)
 XX
 DE His-tag peptide.
 XX
 KW Fibroblast growth factor 20; FGF-20; Parkinson's disease;
 KW substantia nigra; dopaminergic neuron; cochlea-associated disease;
 KW otosclerosis; Cogan's syndrome; Meniere's disease; Pendred's syndrome;
 KW diabetes-associated hearing loss; congenital malformations;
 KW autoimmune disease-related hearing loss; age-related hearing loss;
 KW deafness; ischaemia-related hearing disturbance; immunogen;
 KW antibody; neuro-degenerative disease; tendonitis; wound healing;
 KW stroke; ischaemia; His-tag.
 XX
 OS Synthetic.
 XX
 PN WO200131008-A2.
 XX
 PD 03-MAY-2001.

XX 20-OCT-2000; 2000WO-US29237.
 PF
 XX 22-OCT-1999; 99US-0161162.
 PR
 XX 08-MAR-2000; 2000US-0187856.
 XX
 PA (CHIR) CHIRON CORP.
 PA (KYOU) UNIV KYOTO.
 XX
 PI Itoh N, Kavanaugh WM;
 XX
 XX WPI; 2001-308642/32.
 DR
 XX New human and rat Fibroblast Growth Factor (FGF) polypeptides for
 PT providing trophic support for cells used in Parkinson patients, and
 PT alleviating brain and cochlea conditions -
 PT
 XX Example 6; Page 44; 73pp; English.
 PS
 XX The sequence an His-tag which can be incorporated into a fusion
 CC protein (to allow purification of the fusion protein) with
 CC Human or rat fibroblast growth factor 20, FGF-20. The FGF
 CC polypeptides and nucleic acids encoding them are useful for providing
 CC trophic support for cells in a patient, especially a patient with
 CC Parkinson's disease, and FGF-20 is additionally used to treat patients
 CC with conditions of the substantia nigra. The polypeptides and nucleic
 CC acids are useful for alleviating human brain conditions by slowing
 CC degeneration, restoring function of, or increasing the number of,
 CC dopaminergic neurons. The polypeptides and and nucleic acids are also
 CC useful for alleviating cochlea-associated disease by slowing of the
 CC degeneration of or restoring or maintaining normal function of the
 CC structure of cochlea, where the disease is otosclerosis, Cogan's
 CC syndrome, Meniere's disease, Pendred's syndrome, diabetes-associated
 CC hearing loss, congenital malformations, autoimmune disease-related
 CC hearing loss, age-related hearing loss, deafness associated with lack
 CC of FGF receptor and ischaemia-related hearing disturbance. Other
 CC diseases thought to be amenable to FGF therapeutic activity include
 CC neuro-degenerative diseases, tendonitis, wound healing, stroke and
 CC ischaemia. The polypeptides can be used to screen for agonists and
 CC antagonists. Epitope bearing fragments of the FGF-20 polypeptides can be
 CC used to raise anti-FGF-20 antibodies.
 CC
 XX Sequence 6 AA;
 Query Match 66.7%; Score 4; DB 22; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 HHHH 5
 Db ||||
 1 HHHH 4
 RESULT 73
 AAB97353
 ID AAB97353 standard; peptide; 6 AA.
 XX
 AC AAB97353;
 XX
 DT 15-AUG-2001 (first entry)
 XX
 XX 6-His epitope used in dual labelled receptor construction.
 DE
 XX Automated measurement; cell viability; epitope tag; luminescence;
 KW G-protein coupled receptor; high content screen.
 XX
 OS Synthetic.
 XX
 PN WO200135072-A2.
 XX
 XX 17-MAY-2001.
 PD
 XX 09-NOV-2000; 2000WO-US30896.
 PF

XX 09-NOV-1999; 99US-0164353.
 PR 18-JAN-2000; 2000US-0176504.
 PR (CELL-) CELLOMICS INC.
 PA Ghosh RN, Debiasio R, Chen Y, Bellutta P, Giuliano K, Pasley JW;
 XX WPI; 2001-329169/34.
 XX Automated measurement of cell viability, involves contacting cells with
 PR luminescent reporter molecule, imaging cells to get signals, converting
 PT signals into digital data, and using data to measure viable cell
 PT Example 6; Page 52; 155pp; English.
 XX This invention relates to a method for the automated measurement of cell
 XX viability. The method involves contacting cells with luminescent
 XX reporter molecules, imaging cells to get signals, and converting the
 XX signals into digital data which can be used as a measurement of cell
 XX viability. Included in the invention is a computer readable storage
 XX medium comprising a programme which causes the method of the invention
 XX to be activated. The method is useful for cell state identification in
 XX cells. The method is also useful for drug discovery. An example of the
 XX invention relates to the use of inserted sequences and their ligands for
 XX high content screens incorporating dual labelled receptors. The present
 XX sequence represents an epitope tag used to label one end of a G-protein
 XX coupled receptor (GPCR). The intracellular and extracellular domains of
 XX the GPCR are distinctly labelled so that using the method of the
 XX invention the extent of internalisation of the receptor can be measured.
 XX
 XX Sequence 6 AA;
 SQ

Query Match 66.7%; Score 4; DB 22; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
 DB 1 HHHH 4

RESULT 74
 AAB62662
 D AAB62662 standard; peptide; 6 AA.
 CX AAB62662;
 CX
 CX 23-JUL-2001 (first entry)
 CX C-terminal HIS tag.
 CX Cytokine receptor; zcytor16; IL-TIF; antiinflammatory; cytostatic;
 CX antirheumatic; antiarthritic; antiasthmatic; antiatherosclerotic;
 CX immunosuppressive; chromosome 6q24.1-25.2; human.
 CX Synthetic.
 CX WO200140467-A1.
 CX
 CX 07-JUN-2001.
 CX
 CX 01-DEC-2000; 2000WO-US32703.
 CX
 CX 03-DEC-1999; 99US-0169049.
 CX 13-SEP-2000; 2000US-0232219.
 CX 31-OCT-2000; 2000US-0244610.
 CX (ZYMO) ZYMOGENETICS INC.
 CX Presnell SR, Xu W, Kindsvogel W, Chen Z;
 CX WPI; 2001-356158/37.

XX New soluble cytokine receptor polypeptides and polynucleotides, useful
 PT for diagnosing and treating cancer and inflammatory conditions -
 XX Example 1; Page 193; 210pp; English.
 XX The invention relates to a human cytokine receptor polypeptide,
 CC designated zcytor16. The zcytor16 polypeptide can be expressed by
 CC standard recombinant methodology and can bind to IL-TIF (undefined). The
 CC zcytor16 protein is useful for: inhibiting IL-TIF induced proliferation
 CC or differentiation of hematopoietic cell(s) (progenitors); reducing
 CC IL-TIF induced or IL-9 induced inflammation; and suppressing an
 CC inflammatory response in a mammal with inflammation. Heteromeric/
 CC multimeric receptor polypeptides such as soluble zcytor 16/CRF2-4 can be
 CC used to reduce progression and symptoms of cancer. Zcytor16 polypeptides
 CC can also be used to detect IL-TIF levels which is indicative of
 CC pathological conditions including inflammatory states (e.g. rheumatoid
 CC arthritis) and cancer. Antibodies that bind zcytor16 polypeptides and the
 CC polypeptides themselves are useful for the treatment of inflammation,
 CC inflammatory diseases (e.g. infection, asthma, inflammatory bowel
 CC disease, rheumatoid arthritis and atherosclerosis) and autoimmune
 CC diseases. The antibodies and zcytor16 polynucleotides are also useful
 CC for detecting cancer. The present sequence represents a C-terminal HIS
 CC tag, used in the construction of a zcytor16 mammalian expression
 CC vector.
 XX
 SQ Sequence 6 AA;
 Query Match 66.7%; Score 4; DB 22; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
 DB 1 HHHH 4

RESULT 75
 AAE01997
 ID AAE01997 standard; peptide; 6 AA.
 XX AAE01997;
 AC
 CX 31-JUL-2001 (first entry)
 DT
 XX Poly-His tag used to prepare poly-His version of huRANKL.
 DE
 CX Receptor activator of NF-kappaB; RANK; nuclear factor-kappaB;
 KW NF-kappaB; tumour necrosis factor; TNF; type I transmembrane protein;
 KW TNF receptor-associated factor; TRAF; RANK ligand; RANKL; osteopathic;
 KW inflammatory reaction; bone resorption; gene therapy; immunomodulator;
 KW immune system dysfunction; familial expansile osteolysis; FEO;
 KW early onset Paget's disease of bone; EP; cytostatic; poly-his tag.
 XX Synthetic.
 OS
 XX WO200136637-A1.
 XX
 XX 25-MAY-2001.
 XX
 XX 14-NOV-2000; 2000WO-US31459.
 XX
 XX 17-NOV-1999; 99US-0442029.
 XX (IMMUNEX) IMMUNEX CORP.
 XX Anderson DM, Hughes AE;
 XX WPI; 2001-329222/34.
 XX
 XX New DNA encoding a receptor activator of NF-kappaB polypeptide for the
 PT treatment of Paget's disease and Familial Expansile Osteolysis (FEO) -
 PT

PS Example 15; Page 82; 96pp; English.

XX The present invention relates to a novel receptor, referred to as RANK
 CC (receptor activator of NF (nuclear factor)-kappaB), a member of TNF
 CC (tumour necrosis factor) receptor superfamily. RANK is a Type I
 CC transmembrane protein that interacts with TNF receptor-associated
 CC factors (TRAFs). Triggering of RANK by overexpression or co-expression
 CC of RANK and membrane bound RANK ligand (RANKL) results in upregulation
 CC of the transcription factor NF-kappaB, a ubiquitous transcription factor
 CC that is most extensively utilised in cells of the immune system.
 CC Inhibition of NF-kappaB by RANK antagonists is useful in ameliorating
 CC negative effects of inflammatory reactions, and the effects of excess
 CC bone resorption. The RANK DNAs, proteins and their analogues are useful
 CC for the preparation of pharmaceutical compositions, for infecting target
 CC cells for use in gene therapy applications in diagnosing diseases
 CC associated with RANK, and as targets for use in screening assays. They
 CC may be used in the treatment or diagnosis of immune system dysfunction.
 CC The present invention also encompasses gene therapy methods to correct
 CC gene-activating mutations, associated with e.g. familial expansile
 CC osteolysis (FEO) and early onset Paget's disease of bone (EP). The
 CC present sequence is a synthetic poly-his tag used to prepare soluble,
 CC tagged, poly-his version of human RANKL protein.

XX Sequence 6 AA;

Query Match 66.7%; Score 4; DB 22; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
 DB 1 HHHH 4

RESULT 76

AAU00163

ID AAU00163 standard; Peptide; 6 AA.

AC AAU00163;

XX 16-MAY-2001 (first entry)

DE His Tag for purification of interferon beta fusion proteins.

XX Human; interferon beta; antibody; multiple sclerosis; gene therapy;
 KW viral infection; viral hepatitis; cancer; breast cancer; inflammation;
 KW Crohn's disease; acute myeloid leukaemia; Hodgkin's disease;
 KW ulcerative colitis; immunomodulation; His tag.

OS Synthetic.

XX WO200115736-A2.

PN 08-MAR-2001.

XX 25-AUG-2000; 2000WO-DK00471.

PF 27-AUG-1999; 99DK-0001197.

XX 21-OCT-1999; 99US-0160782.

PR 26-NOV-1999; 99DK-0001691.

PR 07-FEB-2000; 2000DK-0000194.

XX 07-MAR-2000; 2000DK-0000363.

PR 14-APR-2000; 2000DK-0000642.

XX (MAXY-) MAXYGEN APS.

XX Pedersen AH, Schambye HT, Andersen KV, Bornaes C, Rasmussen PB;

XX WPI; 2001-218488/22.

XX A conjugate exhibiting interferon beta activity useful for treating
 PT multiple sclerosis comprises a non-polypeptide group covalently
 PT attached to an interferon beta polypeptide -

XX Disclosure; Page 40; 108pp; English.

XX The sequence is a His tag suitable for purification of Human
 CC interferon beta fusion proteins and to facilitate conjugation to
 CC a non-polypeptide moiety. Conjugates of the invention
 CC exhibiting interferon beta activity comprise at least one first
 CC non-polypeptide group covalently attached to an interferon beta
 CC polypeptide, the amino acid sequence of which differs from wild-type
 CC human interferon beta in at least one introduced and at least one removed
 CC amino acid residue comprising an attachment group for the first
 CC non-polypeptide group. The invention also concerns reducing the
 CC immunogenicity and/or increasing functional in vivo half-life and/or
 CC serum half-life of an interferon beta polypeptide comprising introducing
 CC an amino acid residue constituting an attachment group for a first
 CC non-polypeptide group into a position exposed at the surface of the
 CC protein that does not contain such a group and removing an amino acid
 CC residue constituting an attachment group for a first non-polypeptide
 CC group and subjecting the modified peptide to conjugation with the
 CC non-polypeptide group. The conjugate and a cell culture expressing the
 CC mutated polypeptides are useful in the treatment of disease, especially
 CC multiple sclerosis, and for treating mammals having circulating
 CC antibodies against interferon beta 1a or 1b. DNA encoding the mutated
 CC proteins may be used for gene therapy. The DNA and proteins can also be
 CC used to treat viral infections (e.g. viral hepatitis), cancer (e.g.
 CC breast cancer), inflammation, Crohn's disease, acute myeloid leukaemia,
 CC Hodgkin's disease and ulcerative colitis and for immunomodulation.

XX Sequence 6 AA;

Query Match 66.7%; Score 4; DB 22; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5

DB 1 HHHH 4

RESULT 77

AA335439

ID AAB35439 standard; Peptide; 6 AA.

XX AAB35439;

XX 23-MAY-2001 (first entry)

DE Nascent protein detection method related peptide #1.

XX Nascent protein detection; protein analysis; aminoacylated tRNA;
 KW BODIPY marker; disease diagnosis.

OS Unidentified.

XX WO200114578-A1.

XX 01-MAR-2001.

XX 23-AUG-2000; 2000WO-US23233.

XX 25-AUG-1999; 99US-0382736.

PR 25-AUG-1999; 99US-0382950.

XX (AMBE-) AMBERGEN INC.

XX Rothschild KJ, Gite S, Olejnik J;

XX WPI; 2001-168972/17.

XX Method for detecting nascent proteins by fluorescence comprises
 PT misaminoacylating a tRNA molecule with a marker compound, useful for
 PT detecting mutations in proteins, e.g. cancer -

Example 22; Page 153; 204pp; English.

The present invention describes a method of detecting nascent proteins involving aminoacylating a tRNA molecule with a 4,4-difluoro-4-bora-3A,4A-diaza-s-indacene (BODIPY) marker leading to the production of a misaminoacylated tRNA. This enables the detection, isolation and analysis of nascent proteins using UV without the usual accompanying radioactivity problems. It may be used to detect mutations, for example in cancer, Duchenne muscular dystrophy, adenomatous polyposis coli and colon cancer.

Sequence 6 AA;

Query Match 66.7%; Score 4; DB 22; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2y 2 HHHH 5
1 HHHH 4

RESULT 78

ID AAB73340
ID AAB73340 standard; peptide; 6 AA.

AC AAB73340;

DT 22-MAY-2001 (first entry)

DE Peptide #7, a component of a nickel-binding peptide triad.

DE Ligand-binding array; array-based molecular descriptor device;

KW combinatorial selection; ligand-binding component;

KW supramolecular structure; drug screening; drug discovery;

KW synthetic antibody; immunoassay; pyranosyl-RNA; p-RNA;

QW nickel binding peptide triad.

DS Synthetic.

TH Key Location/Qualifiers

TT Modified-site 1

TT /note= "Conjugated to the tryptamine moiety of p-RNA #72 (AAF76122), GGGG-tryptamine (p-RNA #71) and tryptamine-GAAGG (p-RNA #70)*"

QX WO200113126-A1.

QY 22-FEB-2001.

QZ 28-JUL-2000; 2000WO-US20847.

RA 13-AUG-1999; 99US-0374338.

RA (NANO-) NANOGEN INC.

RA (AVET) AVENTIS RES & TECHNOLOGIES GMBH & CO KG.

RA Heller MJ, Windhab N, Anderson RR, Ackley DE, Nova TS, Hoppe H;

RA Hamon CJ;

RA WPI; 2001-202891/20.

RA New array based microelectronic molecular descriptor device, useful for combinatorial selection of intermolecular ligand binding structures, drug screening and discovery, generation of synthetic antibodies, or for immunoassays -

Example 5; Page 54; 106pp; English.

DE The invention relates to a novel array-based molecular descriptor

DE device for interacting with a specific ligand molecule or structure.

DE The device comprises: (a) an array of electronic microlocations; (b)

DE capture intermolecular ligand-binding components coupled to the

CC microlocations; and (c) supramolecular or supermolecular complexes
CC formed on at least certain of the microlocations. The capture
CC intermolecular ligand binding components include a capture programmable
CC pairing component (e.g., an oligonucleotide) and a ligand binding
CC component (e.g., a peptide), where the capture programmable pairing
CC component is the same at the various microlocations and the ligand
CC binding component is different at the various microlocations. The
CC supramolecular complex formed at the microlocations comprises the
CC capture intermolecular ligand binding component, a first programmable
CC pairing component complementary to at least a portion of the capture
CC programmable pairing component which contains a ligand binding
CC component, and a second programmable pairing component complementary
CC to the capture programmable pairing component and which also contains
CC a ligand binding component, and a ligand molecule. On binding of the
CC specific ligand, a specific supermolecular complex is formed via the
CC pairing of the specific first programmable pairing component, the
CC second programmable pairing component and the capture programmable
CC pairing component. The device is useful for combinatorial selection of
CC intermolecular ligand binding structures, drug screening and discovery,
CC molecular recognition processes, generation of new affinity reagents,
CC generation of synthetic antibodies and for immunoassays. The device
CC and methods are useful for carrying out multi-step and multiplex
CC affinity binding reactions in microscopic formats. The device rapidly
CC carries out higher order selectivity of combinatorially produced
CC intermolecular ligand binding components, supramolecular structures or
CC supramolecular complexes. The device also allows the creation of large
CC libraries of peptides, oligonucleotides, and other interesting
CC combinations of organic or inorganic functional groups. The present
CC device provides the high speed and high selectivity necessary to evaluate
CC and screen vast combinatorial libraries of ligand binding components and
CC the large numbers of subsequent supramolecular structures and complexes
CC that form on the array. To illustrate the invention, the specification
CC refers to the pyranosyl-RNA (p-RNA) sequence #72 (AAF76122), which is
CC an example of a capture programmable pairing component. This is
CC conjugated to a peptide (AAB73340 or AAB73341) via a tryptamine
CC moiety at position 6, and is able to base pair with the p-RNA molecules
CC GGGG-tryptamine (#71) and tryptamine-GAAGG (#70), which are conjugated
CC to peptide AAB73340. The p-RNAs anneal such that the peptides are
CC brought into close proximity with each other to form a ligand-binding
CC triad capable of binding nickel.

Sequence 6 AA;

Query Match 66.7%; Score 4; DB 22; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
1 HHHH 4
1 HHHH 5

RESULT 79

ID AAB66801 standard; Peptide; 6 AA.

AC AAB66801;

DT 27-APR-2001 (first entry)

DE 6-His epitope.

DE Analyte-binding enzyme; analyte analysis; epitope.

OS Unidentified.

QX WO200102600-A2.

QY 11-JAN-2001.

QZ 30-JUN-2000; 2000WO-US18057.

QZ 06-JUL-1999; 99US-0347878.

CC phenotype of a cell. (I) comprises: (a) introducing a molecular library
of randomized candidate nucleic acids into several cells; and (b)
screening the cells for a cell exhibiting an altered phenotype, where the
altered phenotype is due to the presence of a transdominant bioactive
agent. The methods are particularly useful for screening intracellular
transdominant effector peptides and RNA molecules selected inside living
cells from randomized pools. (I) is also useful for introducing random
libraries into cells to screen for bioactive compounds. The methods allow
rapid and highly efficient screening of large numbers of random
oligonucleotides and their corresponding expression products in a single
step. In addition, the methods allow screening in the absence of
significant prior characterization of the cellular defect.

XX Sequence 6 AA;
SQ

Query Match 66.7%; Score 4; DB 22; Length 6;
Best Local Similarity 100.0%; Pred. No. 9,3e+05;
Matches 4; Conservative 0; Mismatch 0; Indels 0; Gaps 0;

QY 2 HHHH 5
||||
Db 1 HHHH 4

RESULT 81
AAB59857
ID AAB59857 standard; peptide; 6 AA.
XX
AAAB59857;
XX
30-MAR-2001 (first entry)
DT XX
DE XX
6-His peptide epitope.
XX
6-His peptide epitope; macromolecule trafficking; endosomal system;
KW membrane receptor internalisation.
XX
Unidentified.
OS
WQ200079241-A2.
PN
XX
28-DEC-2000.
PD
XX
21-JUN-2000; 2000WO-US40260.
PF
XX
21-JUN-1999; 99US-0140143.
PR
XX
12-JUL-1999; 99US-0352171.
PR
XX
11-AUG-1999; 99US-0148360.
PR
XX
13-DEC-1999; 99US-0170313.
XX
(CELL-) CELLOMICS INC.
PA
XX
Rubin RA, Gough AH, Ghosh RN, Giuliano KA, Dunlay RT;
PI
XX
WFI; 2001-091619/10.
DR
XX
Identifying compounds modulating macromolecule trafficking through
PT endosomes, using digital data obtained by converting a luminescent
FT signal from cells contacted with the compound -
XX
XX
Example 6; Page 53; 113pp; English.

CC The present invention relates to an automated method for identifying
compounds that induce or inhibit macromolecule trafficking through an
endosomal system. The method comprises treating cells which possess a
luminescently-tagged macromolecule, with a test compound, and obtaining
luminescent signals from the cells. The signal is converted into digital
data that is used to determine if the test compound has induced or
inhibited the trafficking. The method can also be used to identify the
extent of internalisation of membrane receptors, by fusing a labelled
peptide epitope to the different domains of the receptor e.g. the
extracellular domain and intracellular domain. The present sequence is
one such peptide epitope used in the method of the present invention.

XX SQ Sequence 6 AA;
Query Match 66.7%; Score 4; DB 22; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HHHH 5
DB 1 HHHH 4

RESULT 82
ABG71813
ID ABG71813 standard; Peptide; 6 AA.
XX AC
XX ABG71813;
XX 14-APR-2003 (first entry)
XX HIS tag peptide.
XX
XX RANKL; human receptor activator of NFkappaB; osteoprotegerin; OPG;
KW RANK ligand; osteoclastogenesis; osteoclast inhibitor; gene therapy;
XX osteoporosis; bone disease; HIS tag.
XX OS Synthetic.
XX WO200264782-A2.
XX PN
XX 22-AUG-2002.
XX
XX 08-FEB-2002; 2002WO-DK00090.
XX
XX 09-FEB-2001; 2001DK-0000214.
XX 09-FEB-2001; 2001US-267843P.
XX 23-MAR-2001; 2001DK-0000498.
XX 23-MAR-2001; 2001US-278320P.
XX (MAXY-) MAXYGEN HOLDINGS LTD.
XX
XX Haaning JM, Halkier T;
XX WPI; 2002-691592/74.
XX
XX Novel human receptor activator of NFkappaB (hRANK) or human
PT osteoprotegerin (hOPG) variant polypeptides which bind to RANK ligand
PT (RANKL) with equivalent binding affinity as hRANK or hOPG, useful for
PT treating osteoporosis -
XX
XX Disclosure; Page 32; 129pp; English.

XX
XX This invention relates to a novel polypeptide having an amino acid
CC sequence that is different from and is at least about 70% identical to
CC the amino acid sequence of human receptor activator of NFkappaB (hRANK)
CC or human osteoprotegerin (hOPG), and which has a binding affinity to
CC RANK ligand (RANKL) that is at least as high as the binding affinity of
CC hRANK or hOPG to RANKL, as determined by functional competition assay.
CC The protein of the invention may have osteopathic activity and may act
CC as a RANKL-mediated osteoclastogenesis or RANKL-mediated osteoclast
CC activity inhibitor. The nucleotide sequence shown in the invention
CC may be used in gene therapy. The protein of the invention or fusion
CC proteins comprising this protein are useful as a pharmaceutical, and
CC in the preparation of a medicament for treating or preventing
CC osteoporosis, or other bone diseases or diseases associated with
CC binding of RANKL to the RANK receptor. A host cell containing a vector
CC expressing the protein is useful for producing a polypeptide having
CC binding affinity to RANKL, where the polypeptide comprises at least one
CC N- or O-glycosylation site and the host cell is a eukaryotic host cell
CC capable of in vivo glycosylation, and/or the polypeptide is subjected
CC to conjugation to a non-polypeptide moiety in vitro. The protein of the
CC invention has increased functional in vivo half-life and/or serum
CC half-life compared to hRANK or hOPG and has an improved binding

CC affinity to RANKL compared to the binding affinity of hRANK or hOPG to
CC RANKL, as determined by a functional competition assay. The present
CC sequence represents a peptide tag which may be fused to the protein
CC of the invention to facilitate its purification.

XX
XX SQ Sequence 6 AA;
Query Match 66.7%; Score 4; DB 23; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HHHH 5
DB 1 HHHH 4

RESULT 83
ABP56408
ID ABP56408 standard; Peptide; 6 AA.
XX AC
XX ABP56408;
XX
XX 12-MAR-2003 (first entry)
XX Thioresoxin related linker peptide #3.
XX
XX Thioresoxin; TrxA; fusion protein; allogenic protein; linker.
XX OS Synthetic.
XX CN1189539-A.
XX
XX 05-AUG-1998.
XX
XX 31-JAN-1997; 97CN-0100362.
XX
XX 31-JAN-1997; 97CN-0100362.
XX (BIOL-) BIOLOGICAL ENG INST ACAD MILITARY MEDICI.
XX
XX Cui L, Ma Q;
XX WPI; 2002-733391/80.
XX
XX Soluble expression carrier and its fusion protein of novel thioresoxin
XX
XX Claim 8; Page 7 (Disclosure); 14pp; Chinese.

XX
XX The present invention describes a DNA molecule of a fusion protein which
CC is composed of the DNA sequence similar to encoded thioresoxin in
CC structure and the DNA sequence of an encoded allogenic protein. The
CC soluble expression of the fusion protein is efficient. The expressed
CC fusion protein can be easily purified as several amino acid sequences
CC are inserted at active sites of the thioresoxin gene sequence, and
CC hydrolysed by a specific protease as the junction between the two
CC encoded DNA sequences can be cut by the specific protease, obtaining a
CC recombinant allogenic protein the same as the natural protein in one-class
CC structure. The present sequence represents a thioresoxin related linker
CC peptide which is used in the exemplification of the present invention.

XX
XX SQ Sequence 6 AA;
Query Match 66.7%; Score 4; DB 23; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HHHH 5
DB 1 HHHH 4

RESULT 84

AAE28476
ID AAE28476 standard; peptide; 6 AA.
XX AC AAE28476;
XX
DI 27-DEC-2002 (first entry)
XX
DE 6HIS peptide used for transfection enhancement.
XX
KW Tat region; nucleic acid-binding group; cell transfection system; cancer;
KW gene therapy.
XX
OS Unidentified.
XX
PN US6376248-B1.
XX
PD 23-APR-2002.
XX
PF 16-MAR-1998; 98US-0039780.
XX
PR 14-MAR-1997; 97US-0818200.
XX
PA (LIFE-) LIFE TECHNOLOGIES INC.
XX
PI Hawley-Nelson P, Lan J, Shih P, Jessee JA, Schifferli KP;
PI Gebeyehu G, Ciccarone VC, Evans KL;
XX
DR WPI; 2002-680647/73.
XX
PT New peptide comprising Tat sequence linked to nucleic acid-binding
PT group, useful, e.g. in gene therapy, for improving cell-transfection
PT efficiency -
XX
XX Example 1; Column 171-172; 108pp; English.
XX
CC The invention relates to a peptide comprising Tat sequence linked to
CC nucleic acid-binding group. Peptides of the invention are used as
CC components of a cell transfection system particularly for gene therapy
CC (especially of cancer). The present sequence is a chimeric peptide
CC useful for transfection enhancement. This peptide contains a His tail
CC and a RGD peptide. This sequence is used in the exemplification of
CC the invention.
XX
SQ Sequence 6 AA;
Query Match 66.7%; Score 4; DB 23; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HHHH 5
DB 1 HHHH 4
RESULT 85
AAE28606
ID AAE28606 standard; peptide; 6 AA.
XX AC AAE28606;
XX
DT 27-DEC-2002 (first entry)
XX
DE His peptide tag.
XX
KW Cytokine receptor; Zcytor16; IL-TIF; autoimmune disease; dermatological;
KW inflammatory disease; inflammatory bowel disease; rheumatoid arthritis;
KW aethma; systemic lupus erythematosus; myasthenia gravis; pancreatitis;
KW diabetes; atherosclerosis; glomerulonephritis; gene therapy; cytostatic;
KW immunosuppressive; nephrotropic; allergy; placental health; abortion;
cancer.
XX
OS Synthetic.

PN WO200270655-A2.
XX
PD 12-SEP-2002.
XX
PF 04-MAR-2002; 2002WO-US06267.
XX
PR 02-MAR-2001; 2001US-273035P.
PR 27-MAR-2001; 2001US-279232P.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Presnell SR, Xu W, Kindsvogel W, Chen Z;
XX
DR WPI; 2002-698750/75.
XX
PT New Zcytor16 polypeptide useful for treating autoimmune or inflammatory
PT diseases, e.g. inflammatory bowel disease, rheumatoid arthritis,
PT asthma, atherosclerosis, cancer or diabetes, or in assessing
PT therapeutic aspects of IL-TIF -
XX
XX Example 1; Page 195; 221pp; English.
XX
CC The invention relates to cytokine receptor designated as mouse Zcytor16
CC which can bind and antagonise the IL-TIF. The Zcytor16 polypeptide is
CC useful in modulating the immune system by binding Zcytor16 ligand, and
CC thus, preventing the binding of the ligand with endogenous Zcytor16
CC receptor. It is useful for studying human inflammation or immune
CC function, or for treating autoimmune or inflammatory diseases such as
CC inflammatory bowel disease, rheumatoid arthritis, asthma, systemic
CC lupus erythematosus, myasthenia gravis or allergy, atherosclerosis,
CC cancer, diabetes, glomerulonephritis or pancreatitis, or in assessing
CC therapeutic aspects of IL-TIF, chemical therapeutics, anti-IL-TIF
CC antibodies, anti-Zcytor16 antibodies or Zcytor16 soluble receptors.
CC Zcytor16 DNA and the anti-mouse Zcytor16 antibody are useful as probes
CC in detecting gene expression and gene structure, such as in the
CC diagnosis and/or prevention of spontaneous abortions or in monitoring
CC placental health and function. It is also used in gene therapy. The
CC present sequence is a peptide tag used to construct Zcytor16 mammalian
CC expression vector.
XX
SQ Sequence 6 AA;
Query Match 66.7%; Score 4; DB 23; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HHHH 5
DB 1 HHHH 4
RESULT 86
AAO22927
ID AAO22927 standard; Peptide; 6 AA.
XX AC AAO22927;
XX
DT 12-DEC-2002 (first entry)
XX
DE Human papillomavirus-16 (HPV16) related hexa-s-tag epitope.
XX
KW Virucide; cytostatic; E6; E7 fusion protein; HPV; immunogenic; vaccine;
KW fusion partner; immunogenicity; HPV infection; neoplasm; HPV16;
KW human papillomavirus-16; epitope.
XX
OS Unidentified.
XX
PN EP1243655-A1.
XX
PD 25-SEP-2002.
XX
PF 23-MAR-2001; 2001EP-0107271.
XX

PT Identifying regulator polypeptides which influence target
PT transcriptional regulatory regions, useful for treating cancer,
PT comprises introducing host cells expressing the polypeptide into a
PT library of polynucleotides -
XX
PS Disclosure; Page 101; 224pp; English.

XX The invention discloses a method for identifying polynucleotides encoding
CC a regulator polypeptide, whose expression induces activation of a target
CC transcriptional regulatory region in a host cell. The method comprises
CC providing a population of eukaryotic host cells capable of expressing the
CC polypeptide, introducing into the host cell a library of polynucleotides
CC encoding the polypeptides, permitting expression of the polynucleotides
CC then recovering them from the host cells. The target transcriptional
CC regulatory region is operably associated with a polynucleotide encoding a
CC gene product, the expression of which results in host cell death or cause
CC the host cells to exhibit a pre-determined modified phenotype and where
CC the gene product is expressed upon activation of target transcriptional
CC regulatory region. Each candidate regulator polypeptide comprises a
CC candidate peptide and a molecular scaffold fused to the peptide so that
CC the peptide is displayed on the surface of the candidate regulator
CC polypeptide. The methods are useful in selecting and/or screening
CC regulator molecules, such as polypeptides, which directly or indirectly
CC induce or suppress the transcriptional activation of a target
CC transcriptional regulatory region in a eukaryotic host cell. These
CC regulator molecules may be used (e.g. in gene therapy) for preventing or
CC treating cancers (e.g. breast or ovarian cancer), cardiovascular diseases
CC (e.g. arrhythmia, heart failure, ischaemia), obesity, neurodegenerative
CC diseases (e.g. Alzheimer's disease), bone pathologies, dermatologic
CC diseases (e.g. psoriasis), infections (e.g. viral, bacterial), acquired
CC immunodeficiency syndrome (AIDS), in cosmetic applications and in wound
CC healing. The method is also useful in screening regulator molecules that
CC block antibiotic transport mechanisms, in drug toxicities and drug
CC resistance applications and in improving the performance of existing or
CC developmental drugs. It may also be used in immunobiology, inflammation,
CC allergic response and in biotechnology applications. The sequences
CC presented in ABG92946-ABG93029 are examples of regulator polypeptides.

XX Sequence 6 AA;

Query Match 66.7%; Score 4; DB 23; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0;

QY 2 HHHH 5
|||
Db 1 HHHH 4

RESULT 89
AAE26107
ID AAE26107 standard; peptide; 6 AA.

XX AAE26107;

DT 14-NOV-2002 (first entry)

DE His tag peptide used to generate soluble human RANKL protein.

XX RANK; receptor activator of nuclear factor-kappaB; NF-kB; sepsis;
KW immune response; toxic shock; graft-versus-host reaction; therapy; TRAF;
KW tumour necrosis factor receptor-associated factor; immunosuppressive;
KW antibacterial; antiinflammatory.

XX Synthetic.

XX US2002086827-A1.

XX 04-JUL-2002.

XX 30-MAY-2001; 2001US-0871291.

XX 23-DEC-1996; 96US-059978P.

PR 07-MAR-1997; 97US-077181P.
PR 14-OCT-1997; 97US-064671P.
PR 24-MAY-2000; 2000US-0577800.
PR 22-DEC-1997; 97US-0996139.
PR 17-DEC-1999; 99US-0466496.
XX
PA (IMMV) IMMUNEX CORP.

XX Anderson DM;

XX WPI; 2002-642254/69.

DR A novel RANK (receptor activator of nuclear factor-kappaB (NF-kB))
XX polypeptide, useful for inhibiting activation of NF-kB and for
XX regulating an immune or inflammatory response in an individual -
XX
XX Example 15; Page 40; 49pp; English.

XX The invention relates to novel RANK (receptor activator of nuclear factor
CC -kappaB (NF-kB)) proteins and polynucleotides encoding them. Sequences of
CC the invention are useful for inhibiting activation of NF-kappaB. They are
CC useful for regulating an immune or inflammatory response in an individual
CC at risk for an immune or inflammatory response. Inhibition of NF-kappaB
CC by RANK antagonists is useful in ameliorating negative effects of an
CC inflammatory response that results from triggering of RANK, for e.g.
CC in treating toxic shock or sepsis, graft-versus-host reactions or acute
CC inflammatory reactions. Soluble RANK is useful as an adjunct therapy for
CC diseases characterised by neoplastic cells that express RANK. Soluble
CC forms of the receptor are useful in vitro to screen for agonists or
CC antagonists of RANK activity. The cytoplasmic domain of RANK is useful
CC in developing assays for inhibitors of signal transduction, to screen
CC for molecules that inhibit interaction of RANK with tumour necrosis
CC factor receptor-associated factor (TRAF) 2 or TRAF3. The present
CC sequence is a His tag peptide used to generate a soluble, tagged,
CC poly-His version of human RANK ligand (RANKL) protein. This peptide
CC is used in the exemplification of the invention.

XX Sequence 6 AA;

Query Match 66.7%; Score 4; DB 23; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0;

QY 2 HHHH 5
|||
Db 1 HHHH 4

RESULT 90
ABG32021
ID ABG32021 standard; Peptide; 6 AA.

XX ABG32021;

DT 15-NOV-2002 (first entry)

DE Synthetic protein tag, #1.

XX Tag; single-chain multimeric polypeptide; polyethylene glycol; PEG;
KW granulocyte colony stimulating factor; G-CSF; haematopoietic disorder;
KW radiation therapy; chemotherapy; bone marrow transplantation;
KW acquired immunodeficiency syndrome; AIDS; immunodeficiency disease;
KW leukopenia; acute myeloid leukaemia; half-life; clearance;
KW immunogenicity; bioavailability; single chain G-CSF dimer; HIV;
KW antihuman immunodeficiency virus; haemostatic.

XX Synthetic.

XX WO200236626-A1.

XX 10-MAY-2002.

XX 01-NOV-2001; 2001WO-DK00724.

XX WO200238782-A2.
 XX PD 16-MAY-2002.
 XX PF 13-NOV-2001; 2001WO-EP13125.
 XX PR 13-NOV-2000; 2000DE-1056210.
 XX (ARIM-) ARIMEDES BIOTECHNOLOGY GMBH.
 XX PI Orberger G, Hellmuth K, Wagener C;
 XX WPI; 2002-435853/46.
 XX PT Vector system for preparing recombinant adeno-associated viral
 PT particles, used for high-level expression of heterologous therapeutic
 PT proteins in eukaryotic cells -
 XX PS Disclosure; Page 17; 59pp; German.
 XX CC The invention relates to a viral vector system for preparing recombinant
 CC adeno-associated virus (AAV) particles comprising; at least two plasmid
 CC vectors (ABLS983 and ABLS984) that include the two inverted terminal
 CC repeats (ITR) of AAV and additional sequences and plasmid vectors without
 CC ITRs but containing the rep and cap genes of AAV required for replication
 CC and packaging. The system is useful for producing recombinant AAV for
 CC production of a wide range of therapeutic glycoproteins in eukaryotic
 CC cells. The system provides efficient, large scale production of
 CC heterologous proteins in mammalian cells, without requiring an adenovirus
 CC helper. It is not toxic to host cells and does not cause lysis, so
 CC produced proteins are highly pure. The present sequence is that of a
 CC peptide tag, useful to the invention.
 XX SQ Sequence 6 AA;
 Query Match 66.7%; Score 4; DB 23; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 2 HHHH 5
 Db 1 HHHH 4
 RESULT 93
 AAM49702
 ID AAM49702 standard; peptide; 6 AA.
 AC AAM49702;
 XX 28-MAY-2002 (first entry)
 XX KpnI/SalI His-tag region.
 XX US6 gene; immunodominant; glycoprotein D; antigen; serological testing.
 XX Unidentified.
 XX RU2178806-C2.
 XX 27-JAN-2002.
 XX 21-JAN-2000; 2000RU-0101648.
 XX 21-JAN-2000; 2000RU-0101648.
 XX (VECT-) VECTOR VIROLOGY & BIOTECHN RES CENTRE.
 XX Susloparov MA, Susloparov IM, Plyasunov IV;
 XX WPI; 2002-194324/25.
 XX N-PSDB; ABA99644.

XX Recombinant plasmid DNA phedv1 determining expression of gene US6
 PT fragment of herpes simplex type-1 virus encoding immuno-dominant group
 PT of glycoprotein d(gd)hsv-1 in bacterium Escherichia coli cells -
 XX Disclosure; Column 7; 9pp; Russian.
 XX This invention describes a novel recombinant plasmid DNA constructed in
 CC vitro and containing the herpes simplex type-1 virus (HSV-1) US6 gene
 CC fragment which encodes an immunodominant group of glycoprotein D (gD).
 CC This polypeptide shows the antigenic properties of herpes simplex D (gD).
 CC type-1. Purified recombinant protein can be used as HSV-1 antigen for
 CC the serological testing of HSV-1 in clinical practice. This sequence
 CC represents a His-tag sequence useful to the invention.
 XX SQ Sequence 6 AA;
 Query Match 66.7%; Score 4; DB 23; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 2 HHHH 5
 Db 1 HHHH 4
 RESULT 94
 AAO19983
 ID AAO19983 standard; peptide; 6 AA.
 XX AAO19983;
 AC AAO19983;
 XX 24-MAY-2002 (first entry)
 XX Histidine (His)-tag.
 XX Protein array; cell free system; gridded format; covalent; non-covalent;
 KW ligand; phage display library; cellular protein expression profile;
 KW ribosome display library; post-translational modification; histidine-tag;
 KW his-tag.
 XX Synthetic.
 XX WO200214860-A1.
 XX 21-FEB-2002.
 XX 15-AUG-2001; 2001WO-GB03657.
 XX 15-AUG-2000; 2000GB-0020016.
 XX 01-DEC-2000; 2000GB-0029309.
 XX 16-MAR-2001; 2001GB-0006610.
 XX 07-JUN-2001; 2001GB-0013883.
 XX 14-JUL-2001; 2001GB-0017232.
 XX (DISC-) DISCERNA LTD.
 XX He M, Tausig MJ;
 XX WPI; 2002-241989/29.
 XX Protein arrays useful for identifying interactions of arrayed proteins
 PT with other molecules, in which proteins are produced by in vitro
 PT synthesis using cell-free transcription and translation systems -
 XX Disclosure; Page 9; 61pp; English.
 XX The invention relates to a method for producing a protein array made by
 CC transcription and translation of DNA, or by translation of mRNA, by a
 CC cell free system in vitro to produce individual proteins, domains or
 CC peptides distributed in a gridded format. The peptides of the invention
 CC contain amino acid sequences enabling covalent/non-covalent attachment to
 CC a surface/bead, such that the peptides can be arranged in the grid format

after interaction with suitable ligands or reagents on the surface. The method of the invention is useful for identifying interactions of the arrayed proteins with other molecules, including antibodies, other proteins or domains, peptides, small ligands, cell extracts and nucleic acids. The protein array is also useful for identifying interactions of arrayed proteins with other molecules displayed in a library such as a phage display library or ribosome display library. The protein array is useful for studying cellular protein expression profiles and post-translational modification of cellular proteins. This sequence represents the peptide of a histidine tag used in the invention.

Sequence 6 AA;
Query Match 66.7%; Score 4; DB 23; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y 2 HHHH 5
|||
Db 1 HHHH 4

RESULT 95
AAU76910
ID AAU76910 standard; Peptide; 6 AA.

AAU76910;
05-JUN-2002 (first entry)
C terminal His tag protein.

Z-Cytot II; cytokine receptor; atopy; psoriasis; His tag;
interleukin-T-cell inducible factor; IL-TIF; allergy; asthma;
receptor-modulated apoptosis; Th1; immune response; pancreatitis;
type I diabetes; IDDM; pancreatic cancer; Graves disease; SLE;
inflammatory bowel disease; IBD; Crohn's disease; colon cancer;
intestinal cancer; diverticulosis; autoimmune disease; sepsis;
multiple sclerosis; MS; systemic lupus erythematosus;
myasthenia gravis; rheumatoid arthritis; kidney dysfunction.

Synthetic.
WO200212345-A2.
14-FEB-2002.

08-AUG-2001; 2001WO-US24838.
08-AUG-2000; 2000US-223827P.
01-DEC-2000; 2000US-250876P.

(ZYMO) ZYMOGENETICS INC.

Kindsvogel WR, Topouzis S;
WPI; 2002-217182/27.

New soluble cytokine receptor which binds interleukin-T-cell inducible factor and antagonizes its activity in inflammatory and immune diseases such as cancer, diabetes, asthma, sepsis, psoriasis and autoimmune diseases -

Example 7; Page 99; 117pp; English.

This invention relates to the protein and cDNA sequences of a novel soluble cytokine receptor polypeptide designated zcytorII, which binds interleukin-T-cell inducible factor (IL-TIF) or antagonises IL-TIF activity. The protein of the invention is useful for reducing IL-TIF or IL-9 induced inflammation, and inhibiting IL-TIF-induced proliferation. The protein is also useful for suppressing an immune response in a mammal exposed to an antigen or pathogen. Soluble zcytorII receptor or heterodimeric polypeptide is useful for enhancing the in

vivo killing of target tissues by directly stimulating a zcytorII receptor-modulated apoptotic pathway. IL-TIF is involved in promoting Th1-type immune responses and antagonists of IL-TIF have beneficial use against diseases involving such immune responses. Soluble zcytorII heterodimers are useful as antagonists in inflammatory and immune diseases or conditions such as pancreatitis, type I diabetes (IDDM), pancreatic cancer, Graves disease, inflammatory bowel disease (IBD), Crohn's disease, colon and intestinal cancer, diverticulosis, autoimmune disease (e.g. IDDM, multiple sclerosis (MS), systemic lupus erythematosus (SLE), myasthenia gravis, rheumatoid arthritis and IBD), sepsis, asthma, allergy and other atopic diseases, psoriasis and kidney dysfunction. Soluble zcytorII receptor or heterodimeric receptor polypeptides are useful in vivo or in diagnostic applications to detect IL-TIF expressing cancers in vivo or in tissue samples and to prepare antibodies. ZcytorII serves as a target for MAb therapy of cancer where an antagonising MAb inhibits cancer growth and targets immune-mediated killing. The present sequence represents a C terminal His tag peptide fused to the IgG gamma1/zcytor fusion protein, the His tag is used to facilitate purification of the fusion protein

Sequence 6 AA;

Query Match 66.7%; Score 4; DB 23; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
|||
Db 1 HHHH 4

RESULT 96
AAE18829
ID AAE18829 standard; peptide; 6 AA.

AAE18829;
17-MAY-2002 (first entry)

His tag used in the production of FGF-like polypeptides.

Fibroblast growth factor; FGF-like protein; wound healing; bullosa; epidermolysis; erosive gastritis; inflammatory bowel disease; ulcer; oesophagitis; Crohn's disease; hyaline membrane disease; emphysema; pulmonary fibrosis; hepatic cirrhosis; liver failure; angiogenesis; multiple sclerosis; neurodegenerative disease; lung abnormality; viral hepatitis; respiratory distress syndrome; tumour; skin aging; gene therapy; vaccine; human; His tag.

Homo sapiens.
US2002001825-A1.

03-JAN-2002.

02-APR-2001; 2001US-0822485.

31-MAR-2000; 2000US-0540118.

(ITOH/) ITOH N.

Itoh N;

WPI; 2002-187704/24.

Novel fibroblast growth factor-like polypeptide useful for treating, ameliorating and/or preventing dermal wounds, gastric ulcer, Crohn's disease and pulmonary inflammation -
Example 3; Page 31; 63pp; English.

The invention relates to fibroblast growth factor (FGF)-like polypeptides and nucleic acid molecules encoding such polypeptides.

Sequences of the invention are useful for treating, preventing or ameliorating a medical condition. They are useful for treating dermal wounds, epidermolysis, bullous, male pattern alopecia, gastric ulcer, duodenal ulcer, erosive gastritis, oesophagitis, oesophageal reflux disease, inflammatory bowel disease, Crohn's disease, radiation or chemotherapy-induced gut toxicity, hyaline membrane disease, necrosis of the respiratory epithelium, emphysema, pulmonary inflammation, pulmonary fibrosis, hepatic cirrhosis, toxic insults to the liver, fulminant liver failure, viral hepatitis, mucositis, multiple sclerosis and other neurodegenerative diseases, infantile respiratory distress syndrome, bronchopulmonary dysplasia, acute respiratory distress syndrome or other lung abnormalities, tumours of the eye or the other tissues and organs. FGF-like polypeptides are useful stimulating angiogenesis, promoting wound healing, modulating differentiation of neuronal cells, adipocytes and skeletal muscle cells, preventing or ameliorate skin aging, preventing hair loss, stimulating the growth and differentiation of haematopoietic cells and bone marrow cells and maintaining organs before transplantation and for supporting cultures of primary cells and tissues. Sequences of the invention are also used in gene therapy and as vaccines. The present sequence is His tag used in the production of FGF-like polypeptides.

Query Match 66.7%; Score 4; DB 23; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	2 HHHH 5
db	1 HHHH 4

RESULT 97
AAU75289
ID AAU75289 standard; peptide; 6 AA.

AA
AC AAU75289;

DT 21-MAY-2002 (first entry)

DE His6 tag used in creating array of immobilised antibody fragments.

Protein array device; immobilisation region; IR; protein-capture agent;
protein detection; protein expression pattern; proteomic; tumour cell;
disease-related cell; antibody-based diagnostic test; biochip;
His6 tag.

OS Synthetic.

XX
PN
US6329209-B1.

11-DEC-2001.
PD
XX

PF 14-JUL-1999; 99US-0353555.

PR 14-JUL-1998; 98US-0115455.

PA (ZYOM-) ZYOMYX INC.

PI Wagner P, Nock S, Ault-Riche D, Itin C;

DR WPI; 2002-204455/26.

PT Array of protein-capture agents useful for proteomics and assaying
PT differential gene expression at protein level, has a substrate and
PT array of immobilisation regions having many protein-capture agents on
PT the surface -

PS Example 6: Column 41: 34pp: English.

AA The present invention relates to an array device, comprising a substrate
CC defining a surface, an array of space-apart immobilisation regions (IR)
CC

over the surface, the IR having several protein-capture (PC) agents immobilised on the surface through immobilisation groups chemisorbed or physisorbed to the surface, and, one or more border regions surrounding each IR and separating the IR from one another. The device is useful for simultaneous detection of several proteins which are expression products, or their fragments, of a cell or population of cells in an organism, and for various proteomic applications, including assessing patterns of protein expression and modification in cells. The array of protein-capture agents is useful to compare the protein expression patterns of two cells or population of cells, to assay differential gene expression at the protein level. The array is useful for identifying a protein which is overexpressed in disease-related cells e.g. tumour cells, but not in normal cells. The arrays are also suitable for diagnostic applications. The high density of the antibodies on some arrays enables a large number of different, antibody-based diagnostic tests to be formatted onto a single biochip. The assay facilitates parallel detection and analysis of a large number of proteins in a sample. The present sequence for a His6 tag is used in the method of the present invention.

AA	Sequence	6 AA;
SQ		

Query Match 66.7%; Score 4; DB 23; Length 6;

```

Best Local Similarity 100.0%; Freq. NO. 3.3E+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy	2	HFH	5
Db	1	HFH	4

RESULT 98

ID AAE16561 standard; peptide; 6 AA.

AC AAE16561;

DT 09-APR-2002 (first entry)

Epitope tag #5 fused to yeast cell wall protein AGA2.

AA Yeast cell wall protein; AGA2; T cell receptor; multiple sclerosis;
KW cancer; sepsis; autoimmune disease; arthritis; diabetes.

OS Synthetic.

PN US6331391-B1.

PD 18-DEC-2001.

20-JAN-1998; 98US-0009388.

PR 31-MAY-1996; 96US-018741P.

FR 30-MAY-1977; 3705-0800330.
XX

XX
PA (UNII) UNIV ILLINOIS FOUND.

PI WITTRUP KW, KRAENZ DM, KIEKE M, BODER ET;
XX

DR WPI; 2002-121159/16.
XX

PT Selecting proteins for displayability on surface of yeast cells
PT comprises transforming them with vector encoding protein fused to yeast cell wall protein, useful to select proteins with altered affinity or PT specificity for other molecules -

PS Claim 40; Column 60; 59pp; English.

The present invention relates to a method for selecting proteins for displayability on a yeast cell surface. The method comprises transforming yeast cells with a vector that expresses a test protein fused to a yeast cell wall protein (AGA2), contacting the cells with a label that binds to proteins displayed on the cell wall, and isolating label bound cells.

CC where the test protein is from a variegated population generated by
 CC mutagenesis. The invention is also directed to new processes for
 CC engineering T cell receptor for improved binding properties. Improved
 CC T cell receptor molecules are useful in therapies for cancer, sepsis,
 CC autoimmune diseases such as arthritis, diabetes or multiple sclerosis.
 CC The methods are useful to select proteins with altered affinity, altered
 CC specificity or conditional binding. The present sequence is an epitope
 CC tag fused between protein of interest and yeast cell wall protein AGA2.
 XX Sequence 6 AA;
 SQ

Query Match 66.7%; Score 4; DB 23; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DY 2 HHHH 5
 DB 1 HHHH 4

RESULT 99
 ABU0567
 D AAU0567 standard; Peptide; 6 AA.
 AC AAU0567;
 DT 14-FEB-2002 (first entry)
 DE Purification peptide tag.
 DE Compound library screening; signal transduction; peptide tether;
 GW cellular receptor; reporter molecule.
 CS Synthetic.
 CS US6309842-B1.
 SN 30-OCT-2001.
 XD 24-NOV-1997; 97US-0977378.
 XF 03-DEC-1996; 96US-0758307.
 XJ (GLAX) GLAXO WELLCOME INC.
 XK Dower WJ, Gates CM, Heinkel GL, Lalonde G, Mattheakis LC;
 XI Paddon CJ, Schatz PJ;
 XJ WPI; 2002-048370/06.
 XR Screening compounds having capacity to transduce signal through
 XT cellular receptor by contacting complexes having test compound and
 YT tether susceptible to modification with cell having receptor and
 XT reporter molecule -
 XS Disclosure; Fig 6; 50pp; English.

The invention relates to screening compounds for the capacity to
 transduce a signal through a cellular receptor, by contacting complexes
 having a test compound, a tag recording a step in synthesis of the
 screened compound and a tether susceptible to modification by a reporter
 molecule with cells having a receptor and a DNA fragment encoding the
 reporter. Modification of the tether indicates that the complex contains
 the compound. Supports must be provided, each bearing multiple copies of
 a test compound and a tether, and are contacted with cells in order to
 free a portion of the multiple copies of each of the compounds under test
 from the supports. At least one compound transduces a signal through the
 receptor of a cell causing expression of the reporter molecule, and
 isolating the support having the modified tether, which support bears the
 compound transducing the signal. An array of compounds can also be
 provided on a membrane, which is contacted with a cell, where at least
 one compound transduces a signal through the cell receptor, causing
 expression of the reporter molecule, which is released from the cell, and

CC modifies the membrane at a position proximate to the compound transducing
 CC the signal. The modification of the membrane allows identification and
 CC isolation of the compound transducing the signal. The methods are useful
 CC for the identification of compounds with desired properties, allowing the
 CC isolation of novel pharmaceuticals. This sequence represents a
 CC purification peptide tag of the invention.
 XX Sequence 6 AA;
 SQ

Query Match 66.7%; Score 4; DB 23; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DY 2 HHHH 5
 DB 1 HHHH 4

RESULT 100
 ABU08057
 ID ABU08057 standard; Peptide; 6 AA.
 AC ABU08057;
 DT 13-MAY-2003 (first entry)
 DE Synthetic Histidine-6x tag epitope.
 DE His-6x; epitope; capsaisin receptor; pain; sensory neuron; nociceptor;
 KW heat; low extracellular pH; capsaisin; pepper; vanilloid receptor; VR;
 KW analgesic; gene therapy; urinary incontinence; Histidine-6x tag;
 KW urinary bladder hyperreflexia; neuropathic pain; post-herpetic neuralgia;
 KW diabetic neuropathy; carpal tunnel syndrome; phantom limb pain;
 KW analgesic; uropathic; neuroprotective.
 XX Synthetic.
 OS US6482611-B1.
 PN 19-NOV-2002.
 PD 21-SEP-2000; 2000US-0667422.
 PF 23-SEP-1999; 99US-155617P.
 PR (NEUR-) NEUROGEN CORP.
 PA Cortright D, Krause J;
 XX WPI; 2003-298136/29.
 DR New nucleic acid molecule encoding human capsaisin receptors, useful
 XX for preventing or treating urinary incontinence, urinary bladder
 PT hyperreflexia or neuropathic pain, e.g. post-herpetic neuralgia or
 PT diabetic neuropathy -
 PT Disclosure; Column 39; 31pp; English.

The invention discloses a nucleic acid molecule encoding a human
 capsaisin receptor operatively linked to a nucleic acid sequence
 comprising at least one heterologous regulatory element in the
 appropriate orientation for expression. The sensation of pain is
 detected, and the response mediated, by the peripheral terminals of a
 group of specialised small diameter sensory neurons, termed nociceptors.
 The response to heat, low extracellular pH or capsaisin (the compound
 responsible for the hotness of hot peppers) is characterised by the
 persistent activation of nociceptors. The capsaisin receptor is also
 known as the vanilloid receptor (VR). The nucleic acid molecule is useful
 in diagnostic assays for capsaisin receptor polymorphisms and aberrant
 capsaisin receptor expression levels. The assays are useful for
 identifying individuals that are either particularly susceptible or
 insensitive to the types of pain mediated by the capsaisin receptor,
 and for determining which individuals will benefit from, and which will

CC prove refractory to, treatment with modulators of this receptor. The
 CC nucleic acid may also be used as analgesics, and for preventing or
 CC treating conditions (e.g. gene therapy), such as urinary incontinence,
 CC urinary bladder hyperreflexia and neuropathic pain (e.g. post-herpetic
 CC neuralgia, diabetic neuropathy, carpal tunnel syndrome or phantom limb
 CC pain in amputees). The sequence presented is the synthetic histidine-6x
 CC tag epitope which could be fused to the polypeptide to aid purification
 CC of the human capsaicin receptor.
 XX

SQ Sequence 6 AA;

Query Match 66.7%; Score 4; DB 24; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
 ||||
 Db 1 HHHH 4

Search completed: October 21, 2003, 18:23:54
 Job time : 57.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:31:37 ; Search time 35.6667 Seconds
(without alignments)
28.171 Million cell updates/sec

Title: US-10-057-890A-15
Perfect score: 6
Sequence: 1 GHHHS 6

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 629382 seqs, 167460630 residues

Word size : 0

Total number of hits satisfying chosen parameters: 81216

Minimum DB seq length: 0
Maximum DB seq length: 10

Post-processing: Listing first 150 summaries

- Database : Published Applications AA:*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES									
Result No.	Score	Query Match	Length	DB ID	Description				
1	6	100.0	6	15	US-10-057-890A-15	Sequence 15, Appl			
2	5	83.3	6	15	US-10-179-784-39	Sequence 39, Appl			
3	5	83.3	9	9	US-09-821-984-44	Sequence 44, Appl			
4	5	83.3	9	9	US-09-284-663A-25	Sequence 25, Appl			
5	5	83.3	9	9	US-09-854-280-18	Sequence 18, Appl			
6	5	83.3	9	10	US-09-854-208-17	Sequence 27, Appl			
7	5	83.3	9	12	US-10-203-013-27	Sequence 29, Appl			
8	5	83.3	9	12	US-10-203-013-29	Sequence 73, Appl			
9	5	83.3	10	8	US-08-464-363-73	Sequence 2, Appl			
10	5	83.3	10	10	US-09-981-636-2	Sequence 3, Appl			
11	5	83.3	10	10	US-09-981-636-3	Sequence 31, Appl			
12	5	83.3	10	11	US-09-976-935-31	Sequence 60, Appl			
13	5	83.3	10	15	US-10-104-919-60	Sequence 22, Appl			
14	4	66.7	5	12	US-10-104-307-22	Sequence 176, App			
15	4	66.7	5	12	US-10-016-986-176				

4	66.7	6	9	US-09-280-030-61	Sequence 61, Appl
4	66.7	6	9	US-09-771-956-11	Sequence 11, Appl
4	66.7	6	9	US-09-822-485-30	Sequence 30, Appl
4	66.7	6	9	US-09-760-008A-9	Sequence 9, Appl
4	66.7	6	9	US-09-728-911-12	Sequence 12, Appl
4	66.7	6	9	US-09-809-517A-11	Sequence 11, Appl
4	66.7	6	9	US-09-211-691-14	Sequence 14, Appl
4	66.7	6	9	US-09-808-037-33	Sequence 33, Appl
4	66.7	6	9	US-09-888-615-149	Sequence 149, App
4	66.7	6	9	US-09-875-519A-26	Sequence 26, Appl
4	66.7	6	9	US-09-837-992-44	Sequence 44, Appl
4	66.7	6	9	US-09-871-866-18	Sequence 18, Appl
4	66.7	6	9	US-09-801-968-22	Sequence 22, Appl
4	66.7	6	9	US-09-875-338-93	Sequence 93, Appl
4	66.7	6	10	US-09-988-859-5	Sequence 5, Appl
4	66.7	6	10	US-09-798-584-16	Sequence 16, Appl
4	66.7	6	10	US-09-817-764-9	Sequence 9, Appl
4	66.7	6	10	US-09-801-676A-2	Sequence 2, Appl
4	66.7	6	10	US-09-916-940-99	Sequence 99, Appl
4	66.7	6	10	US-09-780-933-14	Sequence 14, Appl
4	66.7	6	10	US-09-900-530A-15	Sequence 15, Appl
4	66.7	6	10	US-09-973-145-4	Sequence 4, Appl
4	66.7	6	10	US-09-731-558-11	Sequence 11, Appl
4	66.7	6	10	US-09-822-698A-22	Sequence 22, Appl
4	66.7	6	10	US-09-144-886-5	Sequence 5, Appl
4	66.7	6	10	US-09-935-868-1	Sequence 1, Appl
4	66.7	6	10	US-09-877-650-18	Sequence 18, Appl
4	66.7	6	11	US-09-814-604-9	Sequence 9, Appl
4	66.7	6	11	US-09-978-917A-41	Sequence 41, Appl
4	66.7	6	11	US-09-948-391A-43	Sequence 43, Appl
4	66.7	6	11	US-09-989-981A-11	Sequence 11, Appl
4	66.7	6	11	US-09-911-569-106	Sequence 106, App
4	66.7	6	11	US-09-922-226-43	Sequence 43, Appl
4	66.7	6	11	US-09-906-311C-15	Sequence 15, Appl
4	66.7	6	11	US-09-782-587B-12	Sequence 12, Appl
4	66.7	6	11	US-09-802-154-22	Sequence 22, Appl
4	66.7	6	11	US-09-961-400-43	Sequence 43, Appl
4	66.7	6	11	US-09-834-597-32	Sequence 32, Appl
4	66.7	6	12	US-09-874-907A-2	Sequence 2, Appl
4	66.7	6	12	US-09-932-165-1480	Sequence 1480, Ap
4	66.7	6	12	US-10-032-214-299	Sequence 299, App
4	66.7	6	12	US-10-037-243-46	Sequence 46, Appl
4	66.7	6	12	US-10-066-965A-9	Sequence 9, Appl
4	66.7	6	12	US-10-195-707B-35	Sequence 35, Appl
4	66.7	6	12	US-10-200-879-106	Sequence 106, App
4	66.7	6	12	US-10-282-162-1	Sequence 1, Appl
4	66.7	6	12	US-10-190-082-206	Sequence 206, App
4	66.7	6	12	US-10-190-082-217	Sequence 217, App
4	66.7	6	12	US-10-256-705-13	Sequence 13, Appl
4	66.7	6	12	US-10-270-073-20	Sequence 20, Appl
4	66.7	6	12	US-10-291-990-13	Sequence 13, Appl
4	66.7	6	12	US-10-315-964A-400	Sequence 400, App
4	66.7	6	12	US-10-317-251A-400	Sequence 400, App
4	66.7	6	12	US-10-317-252A-400	Sequence 400, App
4	66.7	6	12	US-10-342-103-9	Sequence 9, Appl
4	66.7	6	12	US-10-021-660-135	Sequence 135, App
4	66.7	6	12	US-10-128-587A-63	Sequence 63, Appl
4	66.7	6	12	US-10-340-860A-41	Sequence 41, Appl
4	66.7	6	12	US-09-925-055D-13	Sequence 13, Appl
4	66.7	6	12	US-10-168-080-22	Sequence 22, Appl
4	66.7	6	12	US-10-227-617A-6	Sequence 6, Appl
4	66.7	6	12	US-10-348-966-9	Sequence 9, Appl
4	66.7	6	12	US-09-870-353A-30	Sequence 30, Appl
4	66.7	6	12	US-10-181-804A-13	Sequence 13, Appl
4	66.7	6	12	US-10-328-190-16	Sequence 16, Appl
4	66.7	6	12	US-10-384-743-6	Sequence 6, Appl
4	66.7	6	12	US-09-935-384-780	Sequence 780, App
4	66.7	6	12	US-09-949-039-4	Sequence 4, Appl
4	66.7	6	12	US-10-126-764-3	Sequence 3, Appl
4	66.7	6	12	US-10-190-414-16	Sequence 16, Appl
4	66.7	6	12	US-10-235-175-81	Sequence 81, Appl
4	66.7	6	12	US-09-942-052-709	Sequence 709, App
4	66.7	6	12	US-10-084-706-49	Sequence 49, Appl

89 4 66.7 6 12 US-10-374-207-30 Sequence 30, Appl
90 4 66.7 6 12 US-10-321-558-4 Sequence 4, Appl
91 4 66.7 6 12 US-10-325-717-72 Sequence 72, Appl
92 4 66.7 6 12 US-10-325-720-39 Sequence 39, Appl
93 4 66.7 6 12 US-10-351-189-39 Sequence 39, Appl
94 4 66.7 6 12 US-10-405-878-18 Sequence 18, Appl
95 4 66.7 6 12 US-10-104-307-15 Sequence 15, Appl
96 4 66.7 6 12 US-10-317-773-14 Sequence 14, Appl
97 4 66.7 6 12 US-10-203-531-3 Sequence 3, Appl
98 4 66.7 6 12 US-10-317-428-14 Sequence 14, Appl
99 4 66.7 6 12 US-10-321-204-53 Sequence 53, Appl
100 4 66.7 6 12 US-10-264-127-5 Sequence 5, Appl
101 4 66.7 6 12 US-09-962-756-1778 Sequence 1778, Ap
102 4 66.7 6 14 US-10-003-496-9 Sequence 9, Appl
103 4 66.7 6 14 US-10-005-646-16 Sequence 16, Appl
104 4 66.7 6 14 US-10-158-895-6 Sequence 6, Appl
105 4 66.7 6 14 US-10-029-009-17 Sequence 17, Appl
106 4 66.7 6 14 US-10-093-200A-10 Sequence 10, Appl
107 4 66.7 6 14 US-10-116-273-35 Sequence 35, Appl
108 4 66.7 6 14 US-10-081-395-68 Sequence 68, Appl
109 4 66.7 6 15 US-10-147-087-6 Sequence 6, Appl
110 4 66.7 6 15 US-10-166-232A-6 Sequence 1, Appl
111 4 66.7 6 15 US-10-083-815-1 Sequence 1, Appl
112 4 66.7 6 15 US-10-077-023-93 Sequence 93, Appl
113 4 66.7 6 15 US-10-015-979-12 Sequence 12, Appl
114 4 66.7 6 15 US-10-071-838-13 Sequence 13, Appl
115 4 66.7 6 15 US-10-058-836-6 Sequence 6, Appl
116 4 66.7 6 15 US-10-035-045-24 Sequence 24, Appl
117 4 66.7 6 15 US-10-092-934-11 Sequence 11, Appl
118 4 66.7 6 15 US-10-047-391-12 Sequence 12, Appl
119 4 66.7 6 15 US-10-090-365-12 Sequence 12, Appl
120 4 66.7 6 15 US-10-059-261-286 Sequence 286, App
121 4 66.7 6 15 US-10-039-271-80 Sequence 80, Appl
122 4 66.7 6 15 US-10-001-469-1402 Sequence 1402, Ap
123 4 66.7 6 15 US-10-259-460-2 Sequence 2, Appl
124 4 66.7 6 15 US-10-174-368A-4 Sequence 4, Appl
125 4 66.7 6 15 US-10-222-952A-13 Sequence 13, Appl
126 4 66.7 6 15 US-10-093-248-10 Sequence 10, Appl
127 4 66.7 6 15 US-10-119-235-2 Sequence 2, Appl
128 4 66.7 6 15 US-10-104-919-12 Sequence 12, Appl
129 4 66.7 6 15 US-10-153-862-10 Sequence 10, Appl
130 4 66.7 6 15 US-10-287-035-1 Sequence 1, Appl
131 4 66.7 6 15 US-10-128-590-63 Sequence 63, Appl
132 4 66.7 6 15 US-10-322-142-17 Sequence 17, Appl
133 4 66.7 6 15 US-10-270-071-2 Sequence 2, Appl
134 4 66.7 6 15 US-10-134-643-25 Sequence 25, Appl
135 4 66.7 6 15 US-10-308-515-3 Sequence 3, Appl
136 4 66.7 6 15 US-10-192-294-9 Sequence 9, Appl
137 4 66.7 6 15 US-10-192-294-9 Sequence 9, Appl
138 4 66.7 7 9 US-09-809-517A-12 Sequence 12, Appl
139 4 66.7 7 11 US-09-809-517A-13 Sequence 13, Appl
140 4 66.7 7 12 US-09-963-761B-5 Sequence 5, Appl
141 4 66.7 7 12 US-10-016-986-14 Sequence 14, Appl
142 4 66.7 8 9 US-09-760-009A-10 Sequence 10, Appl
143 4 66.7 8 9 US-09-334-477-11 Sequence 11, Appl
144 4 66.7 8 9 US-09-244-984-2 Sequence 2, Appl
145 4 66.7 8 10 US-09-780-933-15 Sequence 15, Appl
146 4 66.7 8 10 US-09-772-719-22 Sequence 22, Appl
147 4 66.7 8 11 US-09-978-917A-42 Sequence 42, Appl
148 4 66.7 8 11 US-09-967-237-22 Sequence 22, Appl
149 4 66.7 8 11 US-09-782-587B-13 Sequence 13, Appl
150 4 66.7 8 12 US-10-195-707B-36 Sequence 36, Appl
Sequence 10, Appl

ALIGNMENTS

RESULT 1
JS-10-057-890A-15
Sequence 15, Application US/10057890A
Publication No. US20030044901A1
GENERAL INFORMATION:
APPLICANT: Coleman, Timothy

; APPLICANT: Mansfield, Brian
; TITLE OF INVENTION: Scaffold Fusion Polypeptides, Composition for Making the Same, an
; FILE REFERENCE: PFS37
; CURRENT APPLICATION NUMBER: US/10/057,890A
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: 60/265,782
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,858
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 15
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-890A-15

Query Match 100.0%; Score 6; DB 15; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHS 6
Db |||||
1 GHHS 6

RESULT 2
US-10-179-784-39
; Sequence 39, Application US/10179784
; Publication No. US20030036647A1
; GENERAL INFORMATION:
; APPLICANT: Shuman, Stewart
; APPLICANT: Srikantha, Verl
; TITLE OF INVENTION: Pharmacological Targeting of Bacterial DNA Ligase
; FILE REFERENCE: D6468
; CURRENT APPLICATION NUMBER: US/10/179,784
; PRIOR FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: US 60/300,727
; PRIOR FILING DATE: 2001-06-24
; NUMBER OF SEQ ID NOS: 41
; SEQ ID NO 39
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: a histone tag
US-10-179-784-39

Query Match 83.3%; Score 5; DB 15; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHS 5
Db |||||
1 GHHS 5

RESULT 3
US-09-821-984-44
; Sequence 44, Application US/09821984
; Patent No. US20020004205A1
; GENERAL INFORMATION:
; APPLICANT: Consler, Thomas G.
; APPLICANT: Iannone, Marie A.
; APPLICANT: Gray, John G.
; APPLICANT: Stimmel, Julia E.
; TITLE OF INVENTION: METHOD OF INVESTIGATING FUNCTIONAL
; FILE REFERENCE: 07083.000702
; CURRENT APPLICATION NUMBER: US/09/821,984
; CURRENT FILING DATE: 2001-03-30

; PRIOR APPLICATION NUMBER: 60/193,826
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: note =
; OTHER INFORMATION: synthetic construct
US-09-821-984-44

Query Match 83.3%; Score 5; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHH 5
Db 4 GHHH 8

RESULT 4
US-09-284-663A-25
; Sequence 25, Application US/09284663A
; Patent No. US20020012961A1
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Roy, Margaret Ann
; TITLE OF INVENTION: Fibroblast Growth Factor-19
; FILE REFERENCE: P12191(e)
; CURRENT APPLICATION NUMBER: US/09/284,663A
; CURRENT FILING DATE: 1999-04-15
; NUMBER OF SEQ ID NOS: 30
; SEQ ID NO 25
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic epitope-tag.
US-09-284-663A-25

Query Match 83.3%; Score 5; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHH 5
Db 1 GHHH 5

RESULT 5
US-09-854-280-18
; Sequence 18, Application US/09854280
; Patent No. US20020052027A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Jian
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin
; APPLICANT: Li, Hanzhong
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: IL-17 HOMOLOGOUS POLYPEPTIDES AND THERAPEUTIC USES THEREOF
; FILE REFERENCE: P1361R1C2
; CURRENT APPLICATION NUMBER: US/09/854,280
; CURRENT FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: US 09/311,832
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: US 60/085,579

; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: US 60/113,621
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 26
; SEQ ID NO 18
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HIS tag
US-09-854-280-18

Query Match 83.3%; Score 5; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHH 5
Db 1 GHHH 5

RESULT 6
US-09-854-208-18
; Sequence 18, Application US/09854208
; Patent No. US20020106743A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Jian
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin
; APPLICANT: Li, Hanzhong
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: IL-17 HOMOLOGOUS POLYPEPTIDES AND THERAPEUTIC USES
; FILE REFERENCE: P1381-R1
; CURRENT APPLICATION NUMBER: US/09/854,208
; CURRENT FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: US/09/311,832
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: US 60/085,579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: US 60/113,621
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 26
; SEQ ID NO 18
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: Artificial Sequence
; LOCATION: 1-9
; OTHER INFORMATION: His tag
US-09-854-208-18

Query Match 83.3%; Score 5; DB 10; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHH 5
Db 1 GHHH 5

RESULT 7
US-10-203-013-27
; Sequence 27, Application US/10203013
; Publication No. US20030158134A1
; GENERAL INFORMATION:
; APPLICANT: SmithKline Beecham Biologicals S.A.
; TITLE OF INVENTION: No. US20030158134A1el Use
; FILE REFERENCE: B45209
; CURRENT APPLICATION NUMBER: US/10/203,013
; CURRENT FILING DATE: 2002-07-31

PRIOR APPLICATION NUMBER: GB 0002200.4
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: GB 0009336.9
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: GB 0013806.5
PRIOR FILING DATE: 2000-06-06
PRIOR APPLICATION NUMBER: PCT/EP00/05998
PRIOR FILING DATE: 2000-06-28
NUMBER OF SEQ ID NOS: 31
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 27
LENGTH: 9
TYPE: PRT
ORGANISM: Homo sapiens
US-10-203-013-27

Query Match 83.3%; Score 5; DB 12; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHH 5
Db 3 GHHRH 7

RESULT 8
US-10-203-013-29
Sequence 29, Application US/10203013
Publication No. US20030158134A1
GENERAL INFORMATION:
APPLICANT: SmithKline Beecham Biologicals S.A.
TITLE OF INVENTION: No. US20030158134A1e1 Use
FILE REFERENCE: B45209
CURRENT APPLICATION NUMBER: US/10/203,013
CURRENT FILING DATE: 2002-07-31
PRIOR APPLICATION NUMBER: GB 0002200.4
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: GB 0009336.9
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: GB 0013806.5
PRIOR FILING DATE: 2000-06-06
PRIOR APPLICATION NUMBER: PCT/EP00/05998
PRIOR FILING DATE: 2000-06-28
NUMBER OF SEQ ID NOS: 31
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 29
LENGTH: 9
TYPE: PRT
ORGANISM: Homo sapiens
US-10-203-013-29

Query Match 83.3%; Score 5; DB 12; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHH 5
Db 3 GHHRH 7

RESULT 9
US-08-464-363-73
Sequence 73, Application US/08464363
Publication No. US20030035815A1
GENERAL INFORMATION:
APPLICANT: Rogers, Bruce L.
APPLICANT: Morgenstern, Jay
APPLICANT: Bond, Julian F.
APPLICANT: Garman, Richard D.
APPLICANT: Greenstein, Julia L.
APPLICANT: Kuo, Mei-chang
APPLICANT: Morville, Malcolm
TITLE OF INVENTION: RECOMBITOPE PEPTIDES

NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSES: Lahive & Cockfield
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/464,363
FILING DATE: 05-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/807,529
FILING DATE: 13-DEC-1991
APPLICATION NUMBER: US 07/662,276
FILING DATE: 28-FEB-1991
APPLICATION NUMBER: US 07/431,565
FILING DATE: 03-NOV-1989
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-015CN
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
INFORMATION FOR SEQ ID NO: 73:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-464-363-73

Query Match 83.3%; Score 5; DB 8; Length 10;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHH 5
Db 2 GHHRH 6

RESULT 10
US-09-981-636-2
Sequence 2, Application US/09981636
Publication No. US20020182643A1
GENERAL INFORMATION:
APPLICANT: MARKS, JAMES
APPLICANT: NIELSEN, ULRIK
APPLICANT: KIRPOTIN, DMITRI
TITLE OF INVENTION: METHODS OF HIGH-THROUGHPUT SCREENING FOR INTERNALIZING ANTIBODIES
FILE REFERENCE: 407T-897710US
CURRENT APPLICATION NUMBER: US/09/981,636
CURRENT FILING DATE: 2002-06-25
PRIOR APPLICATION NUMBER: US 60/241,279
PRIOR FILING DATE: 2000-10-18
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 10
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: His containing peptide
US-09-981-636-2

Query Match 83.3%; Score 5; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 5.2;

```
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHH 5
Db 4 GHHHH 8

RESULT 11
US-09-981-636-3
; Sequence 3, Application US/09981636
; Publication No. US20020182643A1
; GENERAL INFORMATION:
; APPLICANT: MARKS, JAMES
; APPLICANT: NIELSEN, ULRIC
; APPLICANT: KIRPOTIN, DMITRI
; TITLE OF INVENTION: METHODS OF HIGH-THROUGHPUT SCREENING FOR INTERNALIZING ANTIBODIES
; FILE REFERENCE: 407T-897710US
; CURRENT APPLICATION NUMBER: US/09/981,636
; CURRENT FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/241,279
; PRIOR FILING DATE: 2000-10-18
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: His containing tag
US-09-981-636-3

Query Match 83.3%; Score 5; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHH 5
Db 4 GHHHH 8

RESULT 12
US-09-976-935-31
; Sequence 31, Application US/09976935
; Publication No. US2003008061A1
; GENERAL INFORMATION:
; APPLICANT: Staunton, et al.
; TITLE OF INVENTION: MATERIALS AND METHODS TO MODULATE LIGAND BINDING/ENZYMATIC ACTIVITY
; FILE REFERENCE: 27866/36470A
; CURRENT APPLICATION NUMBER: US/09/976,935
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/239,750
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Histidine tag
US-09-976-935-31

Query Match 83.3%; Score 5; DB 11; Length 10;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHH 5
Db 2 GHHHH 6

RESULT 13
US-10-104-919-60
; Sequence 60, Application US/10104919
```

```
; Publication No. US20030099608A1
; GENERAL INFORMATION:
; APPLICANT: Presnell, Scott R.
; APPLICANT: Xu, Wenfeng
; APPLICANT: Kindsvogel, Wayne
; APPLICANT: Chen, Zhi
; APPLICANT: Hughes, Steven D.
; TITLE OF INVENTION: Human Cytokine Receptor
; FILE REFERENCE: 01-12
; CURRENT APPLICATION NUMBER: US/10/104,919
; CURRENT FILING DATE: 2002-03-23
; PRIOR APPLICATION NUMBER: US 60/279,222
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 60
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C-terminal HIS tag, with spacer
US-10-104-919-60
```

```
Query Match 83.3%; Score 5; DB 15; Length 10;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHH 5
Db 4 GHHHH 8
```

```
RESULT 14
US-10-104-307-22
; Sequence 22, Application US/10104307
; Publication No. US20030180729A1
; GENERAL INFORMATION:
; APPLICANT: GUNNING, Kerry B.
; APPLICANT: POWDRILL, Tom
; APPLICANT: HOGAN, Michael
; TITLE OF INVENTION: Hybridization Rate Enhancement for Substrate-Bound Specific Nu
; FILE REFERENCE: 053960.0001/10S
; CURRENT APPLICATION NUMBER: US/10/104,307
; CURRENT FILING DATE: 2002-03-22
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(5)
; OTHER INFORMATION: synthetic peptide
; NAME/KEY: MOD RES
; LOCATION: (5)..(5)
; OTHER INFORMATION: carboxyl terminal amide cap
US-10-104-307-22
```

```
Query Match 66.7%; Score 4; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
Db 2 HHHH 5
```

```
RESULT 15
US-10-016-986-176
; Sequence 176, Application US/10016986
; Publication No. US20030187247A1
```

```
GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; FILE REFERENCE: 313.2CON1
; CURRENT APPLICATION NUMBER: US/10/016,986
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 09/149,898
; PRIOR FILING DATE: 1998-09-08
; PRIOR APPLICATION NUMBER: US 08/899,575
; PRIOR FILING DATE: 1997-07-24
; PRIOR APPLICATION NUMBER: US 08/276,852
; PRIOR FILING DATE: 1994-07-18
; PRIOR APPLICATION NUMBER: US 08/178,302
; PRIOR FILING DATE: 1994-01-06
; PRIOR APPLICATION NUMBER: PCT/US93/09328
; PRIOR FILING DATE: 1993-09-30
; PRIOR APPLICATION NUMBER: US 07/954,148
; PRIOR FILING DATE: 1992-09-30
; NUMBER OF SEQ ID NOS: 176
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 176
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-10-016-986-176

Query Match          66.7%; Score 4; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 HHHH 5
Db      1 HHHH 4

RESULT 16
US-09-280-030-61
; Sequence 61, Application US/09280030A
; Patent No. US2001002151A1
; GENERAL INFORMATION:
; APPLICANT: Sato, Seiji
; APPLICANT: Higashikuni, Naohiko
; APPLICANT: Kudo, Toshiyuki
; APPLICANT: Kondo, Masaaki
; TITLE OF INVENTION: DNAS ENCODING NEW FUSION PROTEINS AND PROCESSES FOR
; TITLE OF INVENTION: PREPARING USEFUL POLYPEPTIDES THROUGH EXPRESSION OF THE
; TITLE OF INVENTION: DNAS
; FILE REFERENCE: 382.1026
; CURRENT APPLICATION NUMBER: US/09/280,030A
; CURRENT FILING DATE: 1999-03-26
; EARLIER APPLICATION NUMBER: JP10-87339/1998
; EARLIER FILING DATE: 1998-03-31
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Designated is
; OTHER INFORMATION: a tag for separation/purification of a fusion
; OTHER INFORMATION: protein
US-09-280-030-61

Query Match          66.7%; Score 4; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 HHHH 5
Db      1 HHHH 4

GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; FILE REFERENCE: 313.2CON1
; CURRENT APPLICATION NUMBER: US/10/016,986
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 09/149,898
; PRIOR FILING DATE: 1998-09-08
; PRIOR APPLICATION NUMBER: US 08/899,575
; PRIOR FILING DATE: 1997-07-24
; PRIOR APPLICATION NUMBER: US 08/276,852
; PRIOR FILING DATE: 1994-07-18
; PRIOR APPLICATION NUMBER: US 08/178,302
; PRIOR FILING DATE: 1994-01-06
; PRIOR APPLICATION NUMBER: PCT/US93/09328
; PRIOR FILING DATE: 1993-09-30
; PRIOR APPLICATION NUMBER: US 07/954,148
; PRIOR FILING DATE: 1992-09-30
; NUMBER OF SEQ ID NOS: 176
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 176
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-10-016-986-176

Query Match          66.7%; Score 4; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 HHHH 5
Db      1 HHHH 4

RESULT 17
US-09-771-956-11
; Sequence 11, Application US/09771956
; Patent No. US20010031474A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, Michele
; APPLICANT: Brodbeck, Robbin
; APPLICANT: Krause, James
; TITLE OF INVENTION: Chimeric Neuropeptide Y Receptors
; FILE REFERENCE: N2000.001
; CURRENT APPLICATION NUMBER: US/09/771,956
; CURRENT FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:HEXAHISTADINE
; OTHER INFORMATION: TAG
US-09-771-956-11

Query Match          66.7%; Score 4; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 HHHH 5
Db      1 HHHH 4

RESULT 18
US-09-822-485-30
; Sequence 30, Application US/09822485
; Patent No. US20020001825A1
; GENERAL INFORMATION:
; APPLICANT: Itoh, No. US20020001825A1uyuki
; TITLE OF INVENTION: No. US20020001825A1el Fibroblast Growth Factor-Like Polypeptides
; FILE REFERENCE: 08035.0001-01000
; CURRENT APPLICATION NUMBER: US/09/822,485
; CURRENT FILING DATE: 2001-04-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:protein tag
US-09-822-485-30

Query Match          66.7%; Score 4; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 HHHH 5
Db      1 HHHH 4

RESULT 19
US-09-760-008A-9
; Sequence 9, Application US/09760008A
; Patent No. US20020004483A1
; GENERAL INFORMATION:
; APPLICANT: NISSEN, TORBEN LAUESGAARD
; APPLICANT: ANDERSEN, KIM VILBOUR
US-09-760-008A-9
```

APPLICANT: HANSEN, CHRISTIAN KARSTEN
APPLICANT: MIKKELSEN, JAN MOLLER
TITLE OF INVENTION: G-CSF CONJUGATES
FILE REFERENCE: 31-000700US
CURRENT APPLICATION NUMBER: US/09/760,008A
CURRENT FILING DATE: 2001-05-11
PRIOR APPLICATION NUMBER: 60/176,376
PRIOR FILING DATE: 2000-01-14
PRIOR APPLICATION NUMBER: 60/189,506
PRIOR FILING DATE: 2000-03-15
PRIOR APPLICATION NUMBER: 60/215,644
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DK PA 2000 00024
PRIOR FILING DATE: 2000-01-10
PRIOR APPLICATION NUMBER: DK PA 2000 00341
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: DK PA 2000 00943
PRIOR FILING DATE: 2000-06-16
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: peptide
US-09-760-008A-9

Query Match 66.7%; Score 4; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 HHHH 5
|||
Db 1 HHHH 4

RESULT 20
US-09-728-911-12
Sequence 12, Application US/09728911
Patent No. US2002012669A1
GENERAL INFORMATION:
APPLICANT: Presnell, Scott R.
APPLICANT: Xu, Wenfeng
APPLICANT: Kindsvogel, Wayne
APPLICANT: Chen, Zhi
TITLE OF INVENTION: Human Cytokine Receptor
FILE REFERENCE: 99-93
CURRENT APPLICATION NUMBER: US/09/728,911
CURRENT FILING DATE: 2000-12-01
PRIOR APPLICATION NUMBER: US 60/169,049
PRIOR FILING DATE: 1999-12-03
PRIOR APPLICATION NUMBER: US 60/232,219
PRIOR FILING DATE: 2000-09-13
PRIOR APPLICATION NUMBER: US 60/244,610
PRIOR FILING DATE: 2000-10-31
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 12
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: His Tag amino acid sequence
US-09-728-911-12

Query Match 66.7%; Score 4; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 HHHH 5
|||

Db 1 HHHH 4

RESULT 21

US-09-809-517A-11
Sequence 11, Application US/09809517A
Patent No. US20020034733A1
GENERAL INFORMATION:
APPLICANT: Lohning, Corinna
TITLE OF INVENTION: particles via disulfide bonds
FILE REFERENCE: MORPHO/11
CURRENT APPLICATION NUMBER: US/09/809,517A
CURRENT FILING DATE: 2001-03-15
PRIOR APPLICATION NUMBER: EP 99114072.4
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: EP 00103551.8
PRIOR FILING DATE: 2000-02-18
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PatentIn version 3.0
SEQ ID NO 11
LENGTH: 6
TYPE: PRT
ORGANISM: artificial sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic module
US-09-809-517A-11

Query Match 66.7%; Score 4; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 HHHH 5
|||
Db 1 HHHH 4

RESULT 22

US-09-211-691-14
Sequence 14, Application US/09211691
Patent No. US20020034805A1
GENERAL INFORMATION:
APPLICANT: Gilbert, Michel
APPLICANT: Young, N. Martin
APPLICANT: Wakarchuk, Warren W.
APPLICANT: National Research Council of Canada
TITLE OF INVENTION: Fusion Proteins for Use in Enzymatic Synthesis of
FILE REFERENCE: 019957-012910US
CURRENT APPLICATION NUMBER: US/09/211,691
CURRENT FILING DATE: 1998-12-14
PRIOR APPLICATION NUMBER: US 60/069,443
PRIOR FILING DATE: 1997-12-15
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 14
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: 6-His tail for
OTHER INFORMATION: purification
US-09-211-691-14

Query Match 66.7%; Score 4; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 HHHH 5
|||
Db 1 HHHH 4

RESULT 23
US-09-808-037-33
; Sequence 33, Application US/09808037
; Patent No. US20020052311A1
; GENERAL INFORMATION:
; APPLICANT: SOLOMON, Beka
; APPLICANT: HANAN, Eilat
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT AND/OR DIAGNOSIS OF
; TITLE OF INVENTION: NEUROLOGICAL DISEASES AND DISORDERS
; FILE REFERENCE: SOLOMON-2D
; CURRENT APPLICATION NUMBER: US/09/808,037
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 09/629,971
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 09/473,653
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: US 60/152,417
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 33
; LENGTH: 6
; TYPE: PRT
; ORGANISM: synthetic construct
US-09-808-037-33
Query Match 66.7%; Score 4; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HHHH 5
DB 1 HHHH 4
RESULT 24
US-09-888-615-149
; Sequence 149, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: FLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 149
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: His tag
US-09-888-615-149
Query Match 66.7%; Score 4; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
2Y 2 HHHH 5
DB 1 HHHH 4
RESULT 25
US-09-875-519A-26

; Sequence 26, Application US/09875519A
; Patent No. US20020068059A1
; GENERAL INFORMATION:
; APPLICANT: Farries, Timothy C.
; APPLICANT: Harrison, Richard A.
; TITLE OF INVENTION: Down-Regulation Resistant C3 Convertase
; FILE REFERENCE: 4-30443/A/IMU/PCT
; CURRENT APPLICATION NUMBER: US/09/875,519A
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: PCT/GB97/00603
; PRIOR FILING DATE: 1997-03-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: amino acid
; OTHER INFORMATION: insertion
US-09-875-519A-26
Query Match 66.7%; Score 4; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HHHH 5
DB 1 HHHH 4
RESULT 26
US-09-837-992-44
; Sequence 44, Application US/09837992
; Patent No. US20020081687A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Hui
; APPLICANT: Schultz, Joshua
; APPLICANT: Shan, Bei
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: Sickle Cell Disease Susceptibility Gene (SSG): Compositions
; FILE REFERENCE: 018781-006020US
; CURRENT APPLICATION NUMBER: US/09/837,992
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: US 60/198,465
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 60/204,234
; PRIOR FILING DATE: 2000-05-15
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: 6-His epitope
; OTHER INFORMATION: tag
US-09-837-992-44
Query Match 66.7%; Score 4; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HHHH 5
DB 1 HHHH 4
RESULT 27
US-09-871-856-18
; Sequence 18, Application US/09871856
; Patent No. US20020081720A1

GENERAL INFORMATION:
APPLICANT: Anderson, Dirk M.
Galibert, Laurent
Maraskovsky, Eugene
TITLE OF INVENTION: Receptor Activator of NF-kappaB
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation, Law Department
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Power Macintosh 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/871,856
FILING DATE: 31-MAY-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/996,139
FILING DATE: <Unknown>
APPLICATION NUMBER: USSN 08/813,509
FILING DATE: 07 MARCH 1997
APPLICATION NUMBER: USSN 08/772,330
FILING DATE: 23 DECEMBER 1996
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2851-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)233-0644
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 18:

US-09-871-856-18
Query Match 66.7%; Score 4; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
Db 1 HHHH 4

RESULT 28
US-09-801-968-22
Sequence 22, Application US/09801968
Patent No. US20020082205A1
GENERAL INFORMATION:
APPLICANT: Itoh, No. US20020082205A1uyuki
APPLICANT: Kavanaugh, W. Michael
TITLE OF INVENTION: HUMAN PGF-23 GENE AND GENE EXPRESSION
TITLE OF INVENTION: PRODUCTS
FILE REFERENCE: PP-17150.001/201130.40901
CURRENT APPLICATION NUMBER: US/09/801,968
CURRENT FILING DATE: 2001-03-07
NUMBER OF SEQ ID NOS: 46
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 22
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:

OTHER INFORMATION: His6 tag
US-09-801-968-22

Query Match 66.7%; Score 4; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
Db 1 HHHH 4

RESULT 29
US-09-875-338-93
Sequence 93, Application US/09875338
Patent No. US20020095024A1
GENERAL INFORMATION:
APPLICANT: MIKESELL, GLEN E.
APPLICANT: CHANG, HAN
APPLICANT: FINGER, JOSHUA N.
APPLICANT: YANG, GUCHEN
APPLICANT: LU, PIN
APPLICANT: ZHOU, XIA-DI
APPLICANT: PEACH, ROBERT
TITLE OF INVENTION: B7-RELATED NUCLEIC ACIDS AND POLYPEPTIDES USEFUL FOR
FILE OF INVENTION: IMMUNOMODULATION
FILE REFERENCE: 3053-4071US2
CURRENT APPLICATION NUMBER: US/09/875,338
CURRENT FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: 60/272,107
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: 60/209,811
PRIOR FILING DATE: 2000-06-06
NUMBER OF SEQ ID NOS: 94
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 93
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic Epitope tag
US-09-875-338-93

Query Match 66.7%; Score 4; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
Db 1 HHHH 4

RESULT 30
US-09-988-899-5
Sequence 5, Application US/09988899
Patent No. US20020102613A1
GENERAL INFORMATION:
APPLICANT: HOOGENBOOM, HENDRICUS R.J.M.
TITLE OF INVENTION: NOVEL FAB FRAGMENT LIBRARIES AND METHOD FOR THEIR USE
FILE REFERENCE: DX/003 CON
CURRENT APPLICATION NUMBER: US/09/988,899
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: PCT/US00/13682
PRIOR FILING DATE: 2000-05-18
PRIOR APPLICATION NUMBER: 99201558.6
PRIOR FILING DATE: 1999-05-18
NUMBER OF SEQ ID NOS: 71
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: 6X-His tag
US-09-988-899-5

Query Match 66.7%; Score 4; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
|||
DB 1 HHHH 4

RESULT 31
US-09-798-584-16
; Sequence 16, Application US/09798584
; Patent No. US20020102676A1
; GENERAL INFORMATION:
; APPLICANT: Mu, David
; APPLICANT: Powers, Scott
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: KCNE: A No. US20020102676A1el Potassium Channel Protein
; FILE REFERENCE: 018781-004010US
; CURRENT APPLICATION NUMBER: US/09/798,584
; CURRENT FILING DATE: 2001-03-03
; PRIOR APPLICATION NUMBER: US 60/186,951
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: 6-His epitope
US-09-798-584-16

Query Match 66.7%; Score 4; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
|||
DB 1 HHHH 4

RESULT 32
US-09-817-764-9
; Sequence 9, Application US/09817764
; Patent No. US20020102678A1
; GENERAL INFORMATION:
; APPLICANT: HAROCHE, JULIEN
; APPLICANT: ALLIGNET, JEANINE
; APPLICANT: EL SOLH, NEVINE
; TITLE OF INVENTION: STAPHYLOCOCCAL GENE, VGAC, CONFERRING RESISTANCE TO
; FILE REFERENCE: 03495.0201
; CURRENT APPLICATION NUMBER: US/09/817,764
; CURRENT FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/197,372
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic His
US-09-817-764-9

Query Match 66.7%; Score 4; DB 10; Length 6;

Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
|||
DB 1 HHHH 4

RESULT 33
US-09-801-676A-2
; Sequence 2, Application US/09801676A
; Patent No. US20020114734A1
; GENERAL INFORMATION:
; APPLICANT: Pantoliano, Michael W.
; APPLICANT: Bone, Roger F.
; APPLICANT: Rhind, Alexander W.
; APPLICANT: Salemm, Francis R.
; TITLE OF INVENTION: Computer Program for Thermal Shift Assay
; TITLE OF INVENTION: Apparatus for Ligand Development and
; TITLE OF INVENTION: Multi-Variable Protein Chemistry Optimization
; FILE REFERENCE: 1503-011000E
; CURRENT APPLICATION NUMBER: US/09/801,676A
; CURRENT FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 09/459,996
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: 60/017,860
; PRIOR FILING DATE: 1996-05-09
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Six histidine epitope tag of His6D(II)-EGFR1
US-09-801-676A-2

Query Match 66.7%; Score 4; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
|||
DB 1 HHHH 4

RESULT 34
US-09-916-940-99
; Sequence 99, Application US/09916940
; Patent No. US20020127564A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020127564Alan, Garry P
; TITLE OF INVENTION: METHODS FOR SCREENING FOR TRANSDOMINANT INTRACELLULAR
; FILE REFERENCE: A-64260-6/RMS/AMS
; CURRENT APPLICATION NUMBER: US/09/916,940
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 09/727,715
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 08/963,368
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: US 08/589,109
; PRIOR FILING DATE: 1996-01-23
; PRIOR APPLICATION NUMBER: US 08/589,911
; PRIOR FILING DATE: 1996-01-23
; PRIOR APPLICATION NUMBER: US 08/789,333
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 08/787,738
; PRIOR FILING DATE: 1997-01-23
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 99
; LENGTH: 6

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: histidine tag
; OTHER INFORMATION: sequence.
US-09-916-940-99

Query Match 66.7%; Score 4; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
Db 1 HHHH 4

RESULT 35
US-09-780-933-14
; Sequence 14, Application US/09780933
; Patent No. US20020127652A1
; GENERAL INFORMATION:
; APPLICANT: SCHAMBE, HANS T.
; APPLICANT: ANDERSEN, KIM V.
; APPLICANT: VAN DEN HAZEL, BART
; APPLICANT: CHRISTIANSEN, JESPER
; APPLICANT: JEPPESEN, CLAUD B.
; TITLE OF INVENTION: FOLLICLE STIMULATING HORMONES
; FILE REFERENCE: 31-001000US
; CURRENT APPLICATION NUMBER: US/09/780,933
; CURRENT FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: PA 2000 00220
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 60/184,035
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: PA 2000 01092
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,558
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-780-933-14

Query Match 66.7%; Score 4; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
Db 1 HHHH 4

RESULT 36
US-09-900-530A-15
; Sequence 15, Application US/09900530A
; Patent No. US20020128438A1
; GENERAL INFORMATION:
; APPLICANT: Seol, Dae-Wu
; APPLICANT: Billiar, Timothy R.
; TITLE OF INVENTION: DNA Cassette for the Production of
; TITLE OF INVENTION: Secretable Recombinant Trimeric Trail Proteins, Tetracycline
; TITLE OF INVENTION: /Doxycycline-Inducible Adeno-Associated Virus Vector, Their
; TITLE OF INVENTION: Combination and Use in Gene Therapy
; FILE REFERENCE: 5006-1-002
; CURRENT APPLICATION NUMBER: US/09/900,530A
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: KR 2000-38441

; PRIOR FILING DATE: 2000-07-06
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: His-tag
US-09-900-530A-15

Query Match 66.7%; Score 4; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
Db 1 HHHH 4

RESULT 37
US-09-973-145-4
; Sequence 4, Application US/09973145
; Patent No. US20020132248A1
; GENERAL INFORMATION:
; APPLICANT: Rothschild, Kenneth J.
; APPLICANT: Cite, Sadanand
; APPLICANT: Olejnik, Jerzy
; TITLE OF INVENTION: N-Terminal and C-Terminal Markers in Nascent Proteins
; FILE REFERENCE: AMBER-06819
; CURRENT APPLICATION NUMBER: US/09/973,145
; CURRENT FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 09/382,950
; PRIOR FILING DATE: 1999-08-25
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; NAME/KEY: misc feature
; OTHER INFORMATION: Synthetic
US-09-973-145-4

Query Match 66.7%; Score 4; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
Db 1 HHHH 4

RESULT 38
US-09-731-558-11
; Sequence 11, Application US/09731558
; Patent No. US20020146691A1
; GENERAL INFORMATION:
; APPLICANT: Case, Casey Christopher
; APPLICANT: Liu, Qiang
; APPLICANT: Rebar, Edward J.
; APPLICANT: Sangamo Biosciences, Inc.
; TITLE OF INVENTION: Methods of Using Randomized Libraries of Zinc Finger
; TITLE OF INVENTION: Proteins for the Identification of Gene Function
; FILE REFERENCE: 019496-003210US
; CURRENT APPLICATION NUMBER: US/09/731,558
; CURRENT FILING DATE: 2000-12-06
; PRIOR APPLICATION NUMBER: US 09/456,100
; PRIOR FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: 6xHis tag
US-09-731-558-11

Query Match 66.7%; Score 4; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
|
|
|
|
Db 1 HHHH 4

RESULT 39

US-09-822-698A-22
Sequence 22, Application US/09822698A
Patent No. US20020146750A1
GENERAL INFORMATION:

APPLICANT: Hoogenboom, Hendricus R.J.M.
APPLICANT: Henderikx, Maria P.G.
TITLE OF INVENTION: MUCIN-1 Specific Binding Members and Methods of Use Thereof
FILE REFERENCE: DYX-015.1 US
CURRENT APPLICATION NUMBER: US/09/822,698A
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US 09/538,913
PRIOR FILING DATE: 2000-03-30
NUMBER OF SEQ ID NOS: 112
SOFTWARE: Microsoft Word
SEQ ID NO 22
LENGTH: 6
TYPE: PRT
ORGANISM: artificial sequence
FEATURE:
OTHER INFORMATION: hexahistidine peptide

US-09-822-698A-22

Query Match 66.7%; Score 4; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
|
|
|
|
Db 1 HHHH 4

RESULT 40

US-09-144-886-5
Sequence 5, Application US/09144886
Patent No. US20020155114A1
GENERAL INFORMATION:

APPLICANT: Marks, James D
APPLICANT: Amersdorfer, Peter
TITLE OF INVENTION: Therapeutic Monoclonal Antibodies That Neutralize Botulinum Neurotoxins
FILE REFERENCE: 2500.117USO
CURRENT APPLICATION NUMBER: US/09/144,886
CURRENT FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 98
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: His6

US-09-144-886-5

Query Match 66.7%; Score 4; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
|
|
|
|
Db 1 HHHH 4

RESULT 41

US-09-935-868-1
Sequence 1, Application US/09935868
Patent No. US20020164690A1
GENERAL INFORMATION:

APPLICANT: Regeneron Pharmaceuticals, Inc
TITLE OF INVENTION: Receptor Based Antagonists, and Methods of Making and Using
FILE REFERENCE: REG 203D
CURRENT APPLICATION NUMBER: US/09/935,868
CURRENT FILING DATE: 2002-04-11
PRIOR APPLICATION NUMBER: PCT/US99/22045
PRIOR FILING DATE: 1999-09-22
NUMBER OF SEQ ID NOS: 52
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Tag sequence
US-09-935-868-1

Query Match 66.7%; Score 4; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
|
|
|
|
Db 1 HHHH 4

RESULT 42

US-09-877-650-18
Sequence 18, Application US/09877650
Patent No. US20020169117A1
GENERAL INFORMATION:

APPLICANT: Anderson, Dirk M.
Galibert, Laurent
Maraskovsky, Eugene
TITLE OF INVENTION: Ligand for Receptor Activator of NF-kappaB
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation, Law Department
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Power Macintosh 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/877,650
FILING DATE: 08-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/995,659
FILING DATE: 1997-12-22
APPLICATION NUMBER: USSN 08/813,509
FILING DATE: 07 MARCH 1997
APPLICATION NUMBER: USSN 08/772,330
FILING DATE: 23 DECEMBER 1996

Qy 2 HHHH 5
|
|
|
|
Db 1 HHHH 4

ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2852-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)233-0644
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 18:

US-09-877-650-18
Query Match 66.7%; Score 4; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 1 HHHH 4

RESULT 43
US-09-814-604-9
Sequence 9, Application US/09814604
Publication No. US20030003517A1
GENERAL INFORMATION:
APPLICANT: Klein, Elliott S.
APPLICANT: Chandraratna Roshantha A.
TITLE OF INVENTION: Methods of Detecting Dissociated Nuclear
TITLE OF INVENTION: Hormone Receptor Ligands
FILE REFERENCE: P-AR 4528
CURRENT APPLICATION NUMBER: US/09/814,604
CURRENT FILING DATE: 2001-03-22
NUMBER OF SEQ ID NOS: 52
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthetic peptide

US-09-814-604-9
Query Match 66.7%; Score 4; DB 11; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 1 HHHH 4

RESULT 44
US-09-978-917A-41
Sequence 41, Application US/09978917A
Publication No. US20030027299A1
GENERAL INFORMATION:
APPLICANT: Maxygen Aps; Maxygen Holdings
TITLE OF INVENTION: Protein C or activated protein C-like molecules
FILE REFERENCE: 021pus310 - protein C
CURRENT APPLICATION NUMBER: US/09/978,917A
CURRENT FILING DATE: 2001-10-17
NUMBER OF SEQ ID NOS: 48
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 41
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: peptide tag
US-09-978-917A-41

Query Match 66.7%; Score 4; DB 11; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 1 HHHH 4

RESULT 45
US-09-948-391A-43
Sequence 43, Application US/09948391A
Publication No. US20030027311A1
GENERAL INFORMATION:
APPLICANT: Rybak, Susanna M.
APPLICANT: Newton, Dianne L.
APPLICANT: The United States of America
APPLICANT: as represented by The Secretary of the
APPLICANT: Department of Health and Human Services
TITLE OF INVENTION: Recombinant Anti-Tumor RNase
FILE REFERENCE: 015280-343110US
CURRENT APPLICATION NUMBER: US/09/948,391A
CURRENT FILING DATE: 2002-05-10
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: WO PCT/US99/06641
PRIOR FILING DATE: 1999-03-26
PRIOR APPLICATION NUMBER: US 09/622,613
PRIOR FILING DATE: 2000-08-17
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 43
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: six histidine
residue tag at amino terminus

US-09-948-391A-43
Query Match 66.7%; Score 4; DB 11; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 1 HHHH 4

RESULT 46
US-09-989-981A-11
Sequence 11, Application US/09989981A
Publication No. US20030049730A1
GENERAL INFORMATION:
APPLICANT: Hobbs, Helen H.
APPLICANT: Shan, Bei
APPLICANT: Barnes, Robert
APPLICANT: Tian, Hui
APPLICANT: Tularik Inc.
APPLICANT: Board of Regents, The University of Texas System
TITLE OF INVENTION: ABCG5 and ABCG8: Compositions and Methods of Use
FILE REFERENCE: 018781-007320US
CURRENT APPLICATION NUMBER: US/09/989,981A
CURRENT FILING DATE: 2002-07-23
PRIOR APPLICATION NUMBER: US 60/252,235
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/253,645
PRIOR FILING DATE: 2000-11-28
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 11
LENGTH: 6
TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: 6-His epitope
OTHER INFORMATION: tag

US-09-989-981A-11

Query Match 66.7%; Score 4; DB 11; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
Db 1 HHHH 4

RESULT 47
US-09-911-569-106

Sequence 106, Application US/09911569

Publication No. US20030069173A1

GENERAL INFORMATION:

APPLICANT: HAWLEY-NELSON, PAMELA

LAN, JIANQING

SHIH, POJEN

JESSE, JOEL A.

SCHIFFERLI, KEVIN P.

GEBEYEHU, GULILAT

TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSFECTIONS

NUMBER OF SEQUENCES: 120

CORRESPONDENCE ADDRESS:

ADDRESSEE: GREENLEE, WINNER & SULLIVAN

STREET: 5370 MANHATTAN CIRCLE, SUITE 201

CITY: BOULDER

STATE: CO

COUNTRY: US

ZIP: 80303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/911.569

FILING DATE: 23-Jul-2001

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/039,780

FILING DATE: 16-MAR-1998

ATTORNEY/AGENT INFORMATION:

NAME: SULLIVAN, SALLY A.

REGISTRATION NUMBER: 32,064

REFERENCE/DOCKET NUMBER: 32-95D

TELECOMMUNICATION INFORMATION:

TELEPHONE: (303) 499-8080

TELEFAX: (303) 499-8089

INFORMATION FOR SEQ ID NO: 106:

SEQUENCE CHARACTERISTICS:

LENGTH: 6 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

SEQUENCE DESCRIPTION: SEQ ID NO: 106:

US-09-911-569-106

Query Match 66.7%; Score 4; DB 11; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
Db 1 HHHH 4

RESULT 48

US-09-922-226-43

Sequence 43, Application US/09922226

Publication No. US20030077664A1

GENERAL INFORMATION:

APPLICANT: Zhao, Yi

APPLICANT: Thacher, Scott M.

APPLICANT: Xiao, Jia-Hao

APPLICANT: Kusari, Jyotirmoy

APPLICANT: Chandraratna, Roshantha A.

TITLE OF INVENTION: Methods of Screening For Compounds That

Modulate Hormone Receptor Activity

FILE REFERENCE: P-AR 4681

CURRENT APPLICATION NUMBER: US/09/922,226

CURRENT FILING DATE: 2002-01-09

PRIOR APPLICATION NUMBER: US 60/284,797

PRIOR FILING DATE: 2001-04-18

NUMBER OF SEQ ID NOS: 191

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 43

LENGTH: 6

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: synthetic peptide

US-09-922-226-43

Query Match 66.7%; Score 4; DB 11; Length 6;

Best Local Similarity 100.0%; Pred. No. 5.7e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
Db 1 HHHH 4

RESULT 49

US-09-906-311C-15

Sequence 15, Application US/09906311C

Publication No. US20030087827A1

GENERAL INFORMATION:

APPLICANT: Lindberg, Iris

APPLICANT: Cameron, Angus

APPLICANT: Appel, Jon

APPLICANT: Houghten, Richard

TITLE OF INVENTION: Inhibiting Furin With Polybasic Peptides

FILE REFERENCE: 01M36 Lindberg

CURRENT APPLICATION NUMBER: US/09/906,311C

CURRENT FILING DATE: 2001-07-16

NUMBER OF SEQ ID NOS: 31

SOFTWARE: PatentIn version 3.0, WordPerfect version 8

SEQ ID NO 15

LENGTH: 6

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: De novo designed peptide

US-09-906-311C-15

Query Match 66.7%; Score 4; DB 11; Length 6;

Best Local Similarity 100.0%; Pred. No. 5.7e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
Db 1 HHHH 4

RESULT 50

US-09-782-587B-12
; Sequence 12, Application US/09782587B
; Publication No. US20030096338A1
; GENERAL INFORMATION:
; APPLICANT: PEDERSEN, ANDERS H.
; APPLICANT: ANDERSON, KIM V.
; APPLICANT: BORNAES, CLAUS
; TITLE OF INVENTION: FACTOR VII OR VIIA-LIKE MOLECULES
; FILE REFERENCE: 31-001100US
; CURRENT APPLICATION NUMBER: US/09/782,587B
; CURRENT FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: PA 2000 00218
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 60/184,036
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 60/241,916
; PRIOR FILING DATE: 2000-10-18
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic peptide tag

US-09-782-587B-12
Query Match 66.7%; Score 4; DB 11; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 1 HHHH 4

RESULT 51

US-09-802-154-22
; Sequence 22, Application US/09802154
; Publication No. US20030105302A1
; GENERAL INFORMATION:
; APPLICANT: Itoh, No. US20030105302A1uyuki
; APPLICANT: Kavanaugh, W. Michael
; TITLE OF INVENTION: HUMAN FGF-23 GENE AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS
; FILE REFERENCE: PP-17149.001/201130.409
; CURRENT APPLICATION NUMBER: US/09/802,154
; CURRENT FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: His6 tag

US-09-802-154-22
Query Match 66.7%; Score 4; DB 11; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 1 HHHH 4

RESULT 52

US-09-961-400-43
; Sequence 43, Application US/09961400
; Publication No. US20030124131A1

; GENERAL INFORMATION:
; APPLICANT: RYBAK, SUSANNA M.
; APPLICANT: GOLDENBERG, DAVID M.
; APPLICANT: NEWTON, DIANNE L.
; TITLE OF INVENTION: IMMUNOCONJUGATES OF TOXINS DIRECTED AGAINST MALIGNANT
; FILE REFERENCE: 018733/1059
; CURRENT APPLICATION NUMBER: US/09/961,400
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 09/622,613
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: PCT/US99/06641
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/079,751
; PRIOR FILING DATE: 1998-03-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: 6X His tag
US-09-961-400-43

Query Match 66.7%; Score 4; DB 11; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 1 HHHH 4

RESULT 53

US-09-834-597-32
; Sequence 32, Application US/09834597
; Publication No. US20030138925A1
; GENERAL INFORMATION:
; APPLICANT: KEITH, TIM
; TITLE OF INVENTION: NOVEL HUMAN GENE RELATING TO RESPIRATORY DISEASES AND
; TITLE OF INVENTION: OBESITY
; FILE REFERENCE: 2976-4039US
; CURRENT APPLICATION NUMBER: US/09/834,597
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/129,391
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 60/146,336
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 09/548,797
; PRIOR FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: polynhistidine tag

US-09-834-597-32
Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 1 HHHH 4

RESULT 54

US-09-874-907A-2

; Sequence 2, Application US/09874907A
; Publication No. US2003014474A1
; GENERAL INFORMATION:
; APPLICANT: WEIDANZ, JON A.
; APPLICANT: CARD, KIMBERLYN F.
; TITLE OF INVENTION: T CELL RECEPTOR FUSIONS AND CONJUGATES AND METHODS OF
; TITLE OF INVENTION: USE THEREOF
; FILE REFERENCE: 71758/A9890
; CURRENT APPLICATION NUMBER: US/09/874,907A
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: 6X His tag
US-09-874-907A-2

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HHHH 5
DB 1 HHHH 4

RESULT 55

US-09-932-165-1480
; Sequence 1480, Application US/09932165
; Publication No. US20030134784A1
; GENERAL INFORMATION:
; APPLICANT: RAITANO, ARTHUR
; APPLICANT: CHALITA-EID, PIA M.
; APPLICANT: FARIS, MARY
; APPLICANT: SAFFRAN, DOUGLAS
; APPLICANT: AFAR, DANIEL
; APPLICANT: LEVIN, ELANA
; APPLICANT: HUBERT, RENE
; APPLICANT: GE, WANGMAO
; APPLICANT: JAKOBOWITS, AVA
; TITLE OF INVENTION: NUCLEIC ACIDS AND CORRESPONDING PROTEINS ENTITLED
; TITLE OF INVENTION: 83P2H3 AND CAT1P2E11 USEFUL IN TREATMENT AND
; FILE REFERENCE: DETECTION OF CANCER
; FILE REFERENCE: 51158-20014.00
; CURRENT APPLICATION NUMBER: US/09/932,165
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/226,329
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 1508
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1480
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: 6-His tag
US-09-932-165-1480

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HHHH 5
DB 1 HHHH 4

RESULT 56

US-10-032-214-299

; Sequence 299, Application US/10032214
; Publication No. US20030138881A1
; GENERAL INFORMATION:
; APPLICANT: PUNNONEN, JUHA
; APPLICANT: LAZETIC, ALEXANDRA
; APPLICANT: LEONG, STEVEN R.
; APPLICANT: CHANG, CHIA-CHUN
; APPLICANT: APT, DORIS
; APPLICANT: GUSTAFSSON, CLAES
; TITLE OF INVENTION: NOVEL CO-STIMULATORY MOLECULES
; FILE REFERENCE: 02-106730US
; CURRENT APPLICATION NUMBER: US/10/032,214
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: 09/888,324
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: PCT/US01/19973
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,946
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/241,245
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 320
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 299
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Poly-His tag
US-10-032-214-299

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HHHH 5
DB 1 HHHH 4

RESULT 57

US-10-037-243-46
; Sequence 46, Application US/10037243
; Publication No. US20030134352A1
; GENERAL INFORMATION:
; APPLICANT: Brookhaven Science Associates, LLC.
; APPLICANT: Freimuth, Paul I
; APPLICANT: Zhang, Yian-Biao
; APPLICANT: Howitt, Jason A
; TITLE OF INVENTION: Facilitating Protein Folding and Solubility by Use of Peptide Ext
; FILE REFERENCE: BSA 01-22
; CURRENT APPLICATION NUMBER: US/10/037,243
; CURRENT FILING DATE: 2002-06-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Six His tag
US-10-037-243-46

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HHHH 5
DB 1 HHHH 4

RESULT 58

```
US-10-066-965A-9
; Sequence 9, Application US/10066965A
; Publication No. US20030143626A1
; GENERAL INFORMATION:
; APPLICANT: COLAS, PIERRE
; APPLICANT: BRENT, ROGER
; APPLICANT: COHEN, BARAK A.
; TITLE OF INVENTION: TARGETED MODIFICATION OF INTRACELLULAR COMPOUNDS
; FILE REFERENCE: EGYPT 3.0-015
; CURRENT APPLICATION NUMBER: US/10/066,965A
; CURRENT FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: 6-His tag
US-10-066-965A-9
;
; Query Match 66.7%; Score 4; DB 12; Length 6;
; Best Local Similarity 100.0%; Pred. No. 5.7e+05;
; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 2 HHHH 5
; DB 1 HHHH 4
;
; RESULT 59
; US-10-195-707B-35
; Sequence 35, Application US/10195707B
; Publication No. US20030133907A1
; GENERAL INFORMATION:
; APPLICANT: Maxygen Holdings
; TITLE OF INVENTION: Interferon gamma polypeptide variants
; FILE REFERENCE: 23lus510 - INFg variants
; CURRENT APPLICATION NUMBER: US/10/195,707B
; CURRENT FILING DATE: 2002-09-12
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Tag
US-10-195-707B-35
;
; Query Match 66.7%; Score 4; DB 12; Length 6;
; Best Local Similarity 100.0%; Pred. No. 5.7e+05;
; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 2 HHHH 5
; DB 1 HHHH 4
;
; RESULT 60
; US-10-200-879-106
; Sequence 106, Application US/10200879
; Publication No. US20030144230A1
; GENERAL INFORMATION:
; APPLICANT: HAMLEY-NELSON, PAMELA
; APPLICANT: LAN, JUANJING
; APPLICANT: SHIH, POJEN
; APPLICANT: JESSE, JOEL A.
; APPLICANT: SCHIFFERLI, KEVIN P.
; APPLICANT: GEBSEYERU, GULILAT
; TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSFECTIONS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
;
; ADDRESS: GREENLEE, WINNER & SULLIVAN
; STREET: 5370 MANHATTAN CIRCLE, SUITE 201
; CITY: BOULDER
; STATE: CO
; COUNTRY: US
; ZIP: 80303
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/200,879
; FILING DATE: 23-Jul-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/911,569
; FILING DATE: 23-JUL-2001
; APPLICATION NUMBER: US 09/039,780
; FILING DATE: 16-MAR-1998
; APPLICATION NUMBER: US 08/818,200
; FILING DATE: 14-MAR-1997
; APPLICATION NUMBER: US 08/658,130
; FILING DATE: 04-JUN-1996
; APPLICATION NUMBER: US 08/477,354
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: SULLIVAN, SALLY A.
; REGISTRATION NUMBER: 32,064
; REFERENCE/DOCKET NUMBER: 32-95E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303)499-8080
; TELEFAX: (303)499-8089
; INFORMATION FOR SEQ ID NO: 106:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 106:
US-10-200-879-106
;
; Query Match 66.7%; Score 4; DB 12; Length 6;
; Best Local Similarity 100.0%; Pred. No. 5.7e+05;
; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 2 HHHH 5
; DB 1 HHHH 4
;
; RESULT 61
; US-10-282-162-1
; Sequence 1, Application US/10282162
; Publication No. US20030143697A1
; GENERAL INFORMATION:
; APPLICANT: REGENERON PHARMACEUTICALS, INC.
; TITLE OF INVENTION: RECEPTOR BASED ANTAGONISTS, AND METHODS OF MAKING
; TITLE OF INVENTION: AND USING
; FILE REFERENCE: REG 203-B-US
; CURRENT APPLICATION NUMBER: US/10/282,162
; CURRENT FILING DATE: 2002-10-28
; PRIOR FILING DATE: 09/787,835
; PRIOR APPLICATION NUMBER: PCT/US99/22045
; PRIOR FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 6
```

TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Tag sequence
JS-10-282-162-1

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0;

2y 2 HHHH 5
|||
Db 1 HHHH 4

RESULT 62
JS-10-190-082-206
Sequence 206, Application US/10190082
Publication No. US20030148264A1
GENERAL INFORMATION:
APPLICANT: Lasky, Lawrence A.
APPLICANT: Sidhu, Sachdev S.
APPLICANT: Held, Heike A.
TITLE OF INVENTION: PHAGE DISPLAYED PDZ DOMAIN LIGANDS
FILE REFERENCE: P1905R1
CURRENT APPLICATION NUMBER: US/10/190,082
CURRENT FILING DATE: 2002-07-03
PRIOR APPLICATION NUMBER: US 60/303,634
PRIOR FILING DATE: 2001-07-06
NUMBER OF SEQ ID NOS: 683
SEQ ID NO 206
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Synthetic
JS-10-190-082-206

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0;

2y 2 HHHH 5
|||
b 1 HHHH 4

RESULT 63
S-10-190-082-217
Sequence 217, Application US/10190082
Publication No. US20030148264A1
GENERAL INFORMATION:
APPLICANT: Lasky, Lawrence A.
APPLICANT: Sidhu, Sachdev S.
APPLICANT: Held, Heike A.
TITLE OF INVENTION: PHAGE DISPLAYED PDZ DOMAIN LIGANDS
FILE REFERENCE: P1905R1
CURRENT APPLICATION NUMBER: US/10/190,082
CURRENT FILING DATE: 2002-07-03
PRIOR APPLICATION NUMBER: US 60/303,634
PRIOR FILING DATE: 2001-07-06
NUMBER OF SEQ ID NOS: 683
SEQ ID NO 217
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Synthetic
JS-10-190-082-217

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0;

QY 2 HHHH 5
|||
Db 1 HHHH 4

RESULT 64
US-10-256-705-13
Sequence 13, Application US/10256705
Publication No. US20030148330A1
GENERAL INFORMATION:
APPLICANT: Wang, Yan
APPLICANT: MJ Bioworks, Inc.
TITLE OF INVENTION: Improved Nucleic Acid Modifying Enzymes
FILE REFERENCE: 020130-000110US
CURRENT APPLICATION NUMBER: US/10/256,705
CURRENT FILING DATE: 2002-09-27
PRIOR APPLICATION NUMBER: US/09/640,958
PRIOR FILING DATE: 2000-08-16
PRIOR APPLICATION NUMBER: US 60/207,567
PRIOR FILING DATE: 2000-05-26
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 13
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: 6-His epitope
US-10-256-705-13

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
|||
Db 1 HHHH 4

RESULT 65
US-10-270-073-20
Sequence 20, Application US/10270073
Publication No. US20030148409A1
GENERAL INFORMATION:
APPLICANT: ROSSI, EDMUND
APPLICANT: CHANG, CHIEN-HSING KEN
APPLICANT: GOLDENBERG, DAVID
TITLE OF INVENTION: DIRECT TARGETING BINDING PROTEINS
FILE REFERENCE: 042418/0112
CURRENT APPLICATION NUMBER: US/10/270,073
CURRENT FILING DATE: 2002-10-15
PRIOR APPLICATION NUMBER: 60/328,835
PRIOR FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 60/341,881
PRIOR FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 60/345,641
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: 60/404,919
PRIOR FILING DATE: 2002-08-22
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 20
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-270-073-20

Query Match 66.7%; Score 4; DB 12; Length 6;

Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
Db 1 HHHH 4

RESULT 66

US-10-291-990-13
; Sequence 13, Application US/10291990
; Publication No. US20030148457A1
; GENERAL INFORMATION:
; APPLICANT: Bennett Kinrade, Michele
; APPLICANT: Brodbeck, Robbin M.
; APPLICANT: Waters, Stephen
; APPLICANT: Krause, James E.
; TITLE OF INVENTION: Monkey and Canine Melanin Concentrating Hormone Receptors
; FILE REFERENCE: N01.2102
; CURRENT APPLICATION NUMBER: US/10/291,990
; PRIOR FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 60/350,493
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 6
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: His 6x epitope
US-10-291-990-13

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
Db 1 HHHH 4

RESULT 67

US-10-315-964A-400
; Sequence 400, Application US/10315964A
; Publication No. US20030148956A1
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Isfort, Robert J
; APPLICANT: Mazur, Wieslaw A
; TITLE OF INVENTION: Corticotropin Releasing Factor 2 Receptor Agonists
; FILE REFERENCE: 8847M3
; CURRENT APPLICATION NUMBER: US/10/315,964A
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: US 60/349,117
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: US 60/376,337
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: US 60/388,895
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: US 60/411,988
; PRIOR FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 530
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 400
; LENGTH: 6
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Chemically synthesized artificial peptide
US-10-315-964A-400

Query Match 66.7%; Score 4; DB 12; Length 6;

Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
Db 1 HHHH 4

RESULT 68

US-10-317-251A-400
; Sequence 400, Application US/10317251A
; Publication No. US20030148957A1
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Isfort, Robert J
; APPLICANT: Mazur, Wieslaw A
; TITLE OF INVENTION: Corticotropin Releasing Factor 2 Receptor Agonists
; FILE REFERENCE: 8847M2
; CURRENT APPLICATION NUMBER: US/10/317,251A
; CURRENT FILING DATE: 2002-12-11
; PRIOR APPLICATION NUMBER: US 60/349,117
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: US 60/376,337
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: US 60/388,895
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: US 60/411,988
; PRIOR FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 530
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 400
; LENGTH: 6
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Chemically synthesized artificial peptide
US-10-317-251A-400

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
Db 1 HHHH 4

RESULT 69

US-10-317-252A-400
; Sequence 400, Application US/10317252A
; Publication No. US20030148958A1
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Isfort, Robert J
; APPLICANT: Mazur, Wieslaw A
; TITLE OF INVENTION: Corticotropin Releasing Factor 2 Receptor Agonists
; FILE REFERENCE: 8847M
; CURRENT APPLICATION NUMBER: US/10/317,252A
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US 60/349,117
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: US 60/376,337
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: US 60/388,895
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: US 60/411,988
; PRIOR FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 530
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 400
; LENGTH: 6
; TYPE: PRT
; ORGANISM: artificial

FEATURE:
OTHER INFORMATION: Chemically synthesized artificial peptide
US-10-317-252A-400

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 HHHH 5
|
|
|
b 1 HHHH 4

RESULT 70

US-10-342-103-9
Sequence 9, Application US/10342103
Publication No. US20030148359A1
GENERAL INFORMATION:
APPLICANT: Moczydlowski et al.
TITLE OF INVENTION: SAXITOXIN DETECTION AND ASSAY METHOD
FILE REFERENCE: YU-POI-009
CURRENT APPLICATION NUMBER: US/10/342.103
CURRENT FILING DATE: 2003-01-13
PRIOR APPLICATION NUMBER: 60/346086
PRIOR FILING DATE: 2002-01-11
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.2
SEQ ID NO 9
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: His6 tag

US-10-342-103-9

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2 HHHH 5
|
|
|
b 1 HHHH 4

ESULT 71

S-10-021-660-135
Sequence 135, Application US/10021660
Publication No. US20030152926A1
GENERAL INFORMATION:
APPLICANT: Murray, Richard
APPLICANT: Glynn, Richard
APPLICANT: Watson, Susan R.
APPLICANT: EOS Biotechnology, Inc.
TITLE OF INVENTION: NO. US20030152926A1 Methods of Diagnosis of Angiogenesis,
TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
TITLE OF INVENTION: Modulators
FILE REFERENCE: 018501-000710US
CURRENT APPLICATION NUMBER: US/10/021.660
CURRENT FILING DATE: 2001-12-06
PRIOR APPLICATION NUMBER: US/09/784,356
PRIOR FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: US 09/637,977
PRIOR FILING DATE: 2000-08-11
NUMBER OF SEQ ID NOS: 135
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 135
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: epitope tag H1S6
US-10-021-660-135

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
|
|
|
Db 1 HHHH 4

RESULT 72

US-10-128-587A-63
Sequence 63, Application US/10128587A
Publication No. US20030152940A1
GENERAL INFORMATION:
APPLICANT: Inogenetics N.V.
TITLE OF INVENTION: Constructs and methods for expression of recombinant HCV envelope
FILE REFERENCE: 134 PCT
CURRENT APPLICATION NUMBER: US/10/128.587A
CURRENT FILING DATE: 2002-04-24
NUMBER OF SEQ ID NOS: 98
SOFTWARE: PatentIn version 3.1
SEQ ID NO 63
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: adaptor peptide
US-10-128-587A-63

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
|
|
|
Db 1 HHHH 4

RESULT 73

US-10-340-860A-41
Sequence 41, Application US/10340860A
Publication No. US20030152984A1
GENERAL INFORMATION:
APPLICANT: AYGUN, HUSEVIN
APPLICANT: KIRCHER, MARKUS
APPLICANT: ROSMUS, SUSANN
APPLICANT: WOJCZEWSKI, SYLVIA
TITLE OF INVENTION: METHODS FOR THE MANUFACTURE OF DNA
FILE REFERENCE: 0273-0023
CURRENT APPLICATION NUMBER: US/10/340.860A
CURRENT FILING DATE: 2003-01-13
PRIOR APPLICATION NUMBER: EP 02.000720.9
PRIOR FILING DATE: 2002-01-11
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 41
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: 6-His tag
US-10-340-860A-41

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
|
|
|
Db 1 HHHH 4

RESULT 74

US-09-925-055D-13
; Sequence 13, Application US/09925055D
; Publication No. US20030157096A1
; GENERAL INFORMATION:
; APPLICANT: Kindsvogel, Wayne R.
; APPLICANT: Topouzis, Stavros
; TITLE OF INVENTION: SOLUBLE ZCYTOR11 CYTOKINE RECEPTORS
; FILE REFERENCE: 00-56
; CURRENT APPLICATION NUMBER: US/09/925.055D
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/223,827
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: US 60/250,876
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: His tag
US-09-925-055D-13

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0;

QY 2 HHHH 5
Db 1 HHHH 4

RESULT 75

US-10-168-080-22
; Sequence 22, Application US/10168080
; Publication No. US20030157501A1
; GENERAL INFORMATION:
; APPLICANT: SUGIHARA, TAKASHI
; APPLICANT: WADHWA, RENU
; TITLE OF INVENTION: NOVEL HUMAN RNA HELICASE, HELICAIN
; FILE REFERENCE: 084335/0164
; CURRENT APPLICATION NUMBER: US/10/168.080
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/JP00/08908
; PRIOR FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: JP 1999-357406
; PRIOR FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-168-080-22

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0;

QY 2 HHHH 5
Db 1 HHHH 4

RESULT 76

US-10-227-617A-6
; Sequence 6, Application US/10227617A
; Publication No. US20030157084A1

; GENERAL INFORMATION:

; APPLICANT: JAKOBSSON, PER-JOHAN
; APPLICANT: WEINANDER, ROLF
; APPLICANT: MORGENSTERN, RALF
; APPLICANT: LAMBALOT, RALPH H.
; TITLE OF INVENTION: METHODS FOR PREPARING PURIFIED PROSTAGLANDIN E SYNTHASE
; FILE REFERENCE: PFA-005.01
; CURRENT APPLICATION NUMBER: US/10/227.617A
; CURRENT FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: 60/314,853
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/315,405
; PRIOR FILING DATE: 2001-08-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-227-617A-6

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0;

QY 2 HHHH 5
Db 1 HHHH 4

RESULT 77

US-10-318-966-9
; Sequence 9, Application US/10318966
; Publication No. US20030158375A1
; GENERAL INFORMATION:
; APPLICANT: NISSEN, TORBEN LAUESGAARD
; APPLICANT: ANDERSEN, KIM VILBOUR
; APPLICANT: HANSEN, CHRISTIAN KARSTEN
; APPLICANT: MIKKELSEN, JAN MOLLER
; APPLICANT: SCHAMBYE, HANS THALSGAARD
; TITLE OF INVENTION: G-CSF POLYPEPTIDES AND CONJUGATES
; FILE REFERENCE: 0208us620
; CURRENT APPLICATION NUMBER: US/10/318.966
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: 09/904,196
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/760,008
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/176,376
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 60/189,506
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 60/215,644
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DK PA 2000 00024
; PRIOR FILING DATE: 2000-01-10
; PRIOR APPLICATION NUMBER: DK PA 2000 00341
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: DK PA 2000 00943
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-318-966-9

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
|||
Db 1 HHHH 4

RESULT 78
US-09-870-353A-30
; Sequence 30, Application US/09870353A
; Publication No. US20030162173A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Yan
; APPLICANT: Xi, Lei
; APPLICANT: Prosen, Dennis E.
; APPLICANT: MJ Bioworks, Inc.
; TITLE OF INVENTION: Improved Nucleic Acid Modifying Enzymes
; FILE REFERENCE: 020130-000111US
; CURRENT APPLICATION NUMBER: US/09/870,353A
; CURRENT FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: US 60/207,567
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/640,958
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: 6-His epitope tag
JS-09-870-353A-30

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
|||
Db 1 HHHH 4

RESULT 79
JS-10-181-804A-13
; Sequence 13, Application US/10181804A
; Publication No. US20030162220A1
; GENERAL INFORMATION:
; APPLICANT: KIM, EUI HO
; APPLICANT: SUGIYAMA, HARUO
; TITLE OF INVENTION: WT1 INTERACTING PROTEIN WTIP
; FILE REFERENCE: 053466/0335
; CURRENT APPLICATION NUMBER: US/10/181,804A
; CURRENT FILING DATE: 2002-07-27
; PRIOR APPLICATION NUMBER: PCT/JP01/00461
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: JP 2000-14949
; PRIOR FILING DATE: 2000-01-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: 6X His tag
JS-10-181-804A-13

Query Match 66.7%; Score 4; DB 12; Length 6;

Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
|||
Db 1 HHHH 4

RESULT 80
US-10-328-190-16
; Sequence 16, Application US/10328190
; Publication No. US20030162709A1
; GENERAL INFORMATION:
; APPLICANT: ROSSI, EDMUND A.
; APPLICANT: CHANG, CHIEN-HSING KEN
; APPLICANT: GOLDENBERG, DAVID M.
; TITLE OF INVENTION: METHODS OF GENERATING MULTISPECIFIC, MULTIVALENT AGENTS
; FILE REFERENCE: 042418/0115
; CURRENT APPLICATION NUMBER: US/10/328,190
; CURRENT FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: 60/342,103
; PRIOR FILING DATE: 2001-12-26
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: 6-His tag
US-10-328-190-16

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
|||
Db 1 HHHH 4

RESULT 81
US-10-384-743-6
; Sequence 6, Application US/10384743
; Publication No. US20030162228A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/384,743
; CURRENT FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide
US-10-384-743-6

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 1 HHHH 4

RESULT 82
US-09-935-384-780
; Sequence 780, Application US/09935384
; Publication No. US20030166526A1
; GENERAL INFORMATION:
; APPLICANT: CHALLITA-EID, PIA
; APPLICANT: HUBERT, RENE
; APPLICANT: RAITANO, ARTHUR
; APPLICANT: AFAR, DANIEL
; APPLICANT: LEVIN, ELANA
; APPLICANT: FARIS, MARY
; APPLICANT: GE, WANGMAO
; APPLICANT: JAKOBOVITZ, AVA
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN NAMED 158p1H4
; TITLE OF INVENTION: USEFUL IN THE TREATMENT AND DETECTION OF BLADDER AND
; TITLE OF INVENTION: OTHER CANCERS
; FILE REFERENCE: 51158-20033.00
; CURRENT APPLICATION NUMBER: US/09/935.384
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/227,098
; PRIOR FILING DATE: 2000-08-22
; NUMBER OF SEQ ID NOS: 783
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 780
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: 6-His tag
US-09-935-384-780

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 1 HHHH 4

RESULT 83
US-09-949-039-4
; Sequence 4, Application US/09949039
; Publication No. US20030166160A1
; GENERAL INFORMATION:
; APPLICANT: HAWLEY, STEPHEN B.
; TITLE OF INVENTION: COMPOUNDS AND MOLECULAR COMPLEXES COMPRISING MULTIPLE
; TITLE OF INVENTION: BINDING REGIONS DIRECTED TO TRANSCYTOTIC LIGANDS
; FILE REFERENCE: 057220/1301
; CURRENT APPLICATION NUMBER: US/09/949.039
; PRIOR FILING DATE: 2001-09-06
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: 6X His tag
US-09-949-039-4

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5

DB 1 HHHH 4

RESULT 84
US-10-126-764-3
; Sequence 3, Application US/10126764
; Publication No. US20030166834A1
; GENERAL INFORMATION:
; APPLICANT: Bennett Kinrade, Michele
; APPLICANT: Brodbeck, Robbin
; APPLICANT: Krause, James
; TITLE OF INVENTION: MELANIN CONCENTRATING HORMONE RECEPTORS
; FILE REFERENCE: N00.2102
; CURRENT APPLICATION NUMBER: US/10/126,764
; CURRENT FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: 60/284,835
; PRIOR FILING DATE: 2001-04-19
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: His 6x epitope
US-10-126-764-3

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
DB 1 HHHH 4

RESULT 85
US-10-190-414-16
; Sequence 16, Application US/10190414
; Publication No. US20030165996A1
; GENERAL INFORMATION:
; APPLICANT: HALKIER, TORBEN
; APPLICANT: PEDERSEN, ANDERS HJELHOLT
; APPLICANT: OKKELS, JENS SIGURD
; APPLICANT: ANDERSEN, KIM VILBOUR
; TITLE OF INVENTION: METHOD FOR PREPARING MODIFIED POLYPEPTIDES
; FILE REFERENCE: 31-105900US
; CURRENT APPLICATION NUMBER: US/10/190,414
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US/09/611,234A
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 60/189,503
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 60/160,693
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: 60/207,793
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide tag
US-10-190-414-16

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

; OTHER INFORMATION: Description of Artificial Sequence: 6-His tag
; US-09-942-052-709

Query Match      66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred.No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
   ||||
Db 1 HHHH 4

RESULT 86
US-10-235-175-81
; Sequence 81, Application US/10235175
; Publication No. US20030168287A1
; GENERAL INFORMATION:
; APPLICANT: Es van, Helmut
; APPLICANT: Havena, Menzo
; APPLICANT: Verlinden, Stefan
; TITLE OF INVENTION: TARGETED DELIVERY THROUGH A CATIONIC AMINO ACID
; TITLE OF INVENTION: TRANSPORTER
; FILE REFERENCE: 2183-4080US
; CURRENT APPLICATION NUMBER: US/10/235.175
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US/09/315,926
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: EP 99201593.3
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: EP 98201693.3
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 81
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Description of Artificial Sequence: phage
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(6)
; OTHER INFORMATION: /note="HCAT1 amino acid sequence"
; US-10-235-175-81

Query Match      66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred.No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
   ||||
b 1 HHHH 4

RESULT 87
US-09-942-052-709
; Sequence 709, Application US/09942052
; Publication No. US20030170626A1
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Afar, Daniel
; APPLICANT: Ge, Wangmao
; APPLICANT: Chailita-Eid, Pia M.
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3
; TITLE OF INVENTION: USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51159-20028.00
; CURRENT APPLICATION NUMBER: US/09/942.052
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 60/228,432
; PRIOR FILING DATE: 2000-08-28
; NUMBER OF SEQ ID NOS: 744
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 709
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

; OTHER INFORMATION: peptide tag
; US-10-084-706-49

Query Match      66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred.No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
   ||||
Db 1 HHHH 4

RESULT 89
US-10-374-207-30
; Sequence 30, Application US/10374207
; Publication No. US20030170822A1
; GENERAL INFORMATION:
; APPLICANT: Itoh, No. US20030170822A1uyuki
; TITLE OF INVENTION: Fibroblast Growth Factor-Like Molecules and Uses Thereof
; FILE REFERENCE: 08035.0001-02000
; CURRENT APPLICATION NUMBER: US/10/374.207
; CURRENT FILING DATE: 2003-02-25
```

; PRIOR APPLICATION NUMBER: US 09/822,485
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: US 09/540,118
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:protein tag
US-10-374-207-30

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
Db 1 HHHH 4

RESULT 90
US-10-321-558-4
; Sequence 4, Application US/10321558
; Publication No. US20030176663A1
; GENERAL INFORMATION:
; APPLICANT: NERI, DARIO
; APPLICANT: TARLI, LORENZO
; APPLICANT: VITI, FRANCESCA
; APPLICANT: BIRCHLER, MANFRED
; TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTIGRAPHY
; FILE REFERENCE: NOTAR-1 C1
; CURRENT APPLICATION NUMBER: US/10/321,558
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: 09/512,082
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 09/300,425
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 09/075,338
; PRIOR FILING DATE: 1998-05-11
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-321-558-4

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
Db 1 HHHH 4

RESULT 91
US-10-325-717-72
; Sequence 72, Application US/10325717
; Publication No. US20030176328A1
; GENERAL INFORMATION:
; APPLICANT: Rasmussen, Poul Baad
; APPLICANT: Andersen, Kim Vilbourn
; APPLICANT: Pedersen, Anders Hjelholt
; APPLICANT: Schambye, Hans Thalsgaard
; APPLICANT: Halkier, Torben
; APPLICANT: Bogsnes, Are

; TITLE OF INVENTION: Adiponectin Fragments and Conjugates
; FILE REFERENCE: 0251us610
; CURRENT APPLICATION NUMBER: US/10/325,717
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US 60/412,169
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US 60/394,117
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 60/375,492
; PRIOR FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: US 60/343,482
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide tag
US-10-125-717-72

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
Db 1 HHHH 4

RESULT 92
US-10-325-720-39
; Sequence 39, Application US/10325720
; Publication No. US20030175240A1
; GENERAL INFORMATION:
; APPLICANT: Pedersen, A.H., et al.
; APPLICANT: Maxygen Aps
; TITLE OF INVENTION: Interferon-Beta Variants and Conjugates
; FILE REFERENCE: 0202us820
; CURRENT APPLICATION NUMBER: US/10/325,720
; CURRENT FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US 09/648,569
; PRIOR FILING DATE: 2000-08-25
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: peptide tag
US-10-325-720-39

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHH 5
Db 1 HHHH 4

RESULT 93
US-10-351-189-39
; Sequence 39, Application US/10351189
; Publication No. US20030175241A1
; GENERAL INFORMATION:
; APPLICANT: Pedersen, A.H., et al.
; APPLICANT: Maxygen Aps
; TITLE OF INVENTION: Interferon-Beta Variants and Conjugates
; FILE REFERENCE: 0202us830
; CURRENT APPLICATION NUMBER: US/10/351,189

```
CURRENT FILING DATE: 2003-01-24
PRIOR APPLICATION NUMBER: US 09/648,569
PRIOR FILING DATE: 2000-08-25
NUMBER OF SEQ ID NOS: 45
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 39
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: peptide tag
US-10-351-189-39

Query Match      66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred.No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 HHHH 5
      ||||
Db      1 HHHH 4

RESULT 95
US-10-104-307-15
; Sequence 15, Application US/10104307
; Publication No. US20030180729A1
; GENERAL INFORMATION:
; APPLICANT: GUNNING, Kerry B.
; APPLICANT: HOGAN, Michael
; TITLE OF INVENTION: Hybridization Rate Enhancement for Substrate-Bound Specific Nucle
; FILE REFERENCE: 053960.0001/JUS
; CURRENT APPLICATION NUMBER: US/10/104,307
; CURRENT FILING DATE: 2002-03-22
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(6)
; OTHER INFORMATION: synthetic peptide
; NAME/KEY: MOD.RES
; LOCATION: (6)..(6)
; OTHER INFORMATION: carboxyl terminal amide cap
US-10-104-307-15

Query Match      66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred.No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 HHHH 5
      ||||
Db      2 HHHH 5

RESULT 96
US-10-317-773-14
; Sequence 14, Application US/10317773
; Publication No. US20030180928A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Young, N. Martin
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: Fusion Protein Comprising a UDP-GalNAC 4' Epimerase and a
; FILE REFERENCE: 019633-000812US
; CURRENT APPLICATION NUMBER: US/10/317,773
; CURRENT FILING DATE: 2002-12-11
; PRIOR APPLICATION NUMBER: US 60/069,443
; PRIOR FILING DATE: 1997-12-15
; PRIOR APPLICATION NUMBER: US 09/211,691
; PRIOR FILING DATE: 1998-12-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:6-His tail for
; OTHER INFORMATION: purification

US-10-405-878-18
Sequence 18, Application US/10405878
Publication No. US20030175840A1
GENERAL INFORMATION:
APPLICANT: Anderson, Dirk M.
Galibert, Laurent
Maraskovsky, Eugene
TITLE OF INVENTION: Receptor Activator of NF-kappaB
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation, Law Department
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Power Macintosh 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/405,878
FILING DATE: 01-Apr-2003
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/996,139
FILING DATE: 22 DECEMBER 1997
APPLICATION NUMBER: USSN 60/064,671
FILING DATE: 14 OCTOBER 1997
APPLICATION NUMBER: USSN 08/813,509
FILING DATE: 07 MARCH 1997
APPLICATION NUMBER: USSN 08/772,330
FILING DATE: 23 DECEMBER 1996
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2851-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)233-0644
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-10-405-878-18
```

```
US-10-317-773-14
Query Match      66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
Db 1 HHHH 4

RESULT 97
US-10-203-531-3
; Sequence 3, Application US/10203531
; Publication No. US20030186386A1
; GENERAL INFORMATION:
; APPLICANT: Hansen, Christian Karsten
; APPLICANT: Andersen, Kim Vilbour
; TITLE OF INVENTION: Improved Interleukin-10
; FILE REFERENCE: 0215us310
; CURRENT APPLICATION NUMBER: US/10/203,531
; PRIOR FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: DK PA 2000 00221
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/184,024
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: DK PA 2000 01325
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/230,093
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-203-531-3

Query Match      66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
Db 1 HHHH 4

US-10-317-428-14
; Sequence 14, Application US/10317428
; Publication No. US20030186414A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Young, N. Martin
; APPLICANT: Waxarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Nucleic Acid That Encodes a Fusion Protein
; FILE REFERENCE: 019633-000811US
; CURRENT APPLICATION NUMBER: US/10/317,428
; CURRENT FILING DATE: 2002-12-11
; PRIOR APPLICATION NUMBER: US 60/069,443
; PRIOR FILING DATE: 1997-12-15
; PRIOR APPLICATION NUMBER: US 09/211,691
; PRIOR FILING DATE: 1998-12-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

US-10-317-428-14
Query Match      66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
Db 1 HHHH 4

US-10-264-127-5
; Sequence 5, Application US/10264127
; Publication No. US20030190643A1
; GENERAL INFORMATION:
; APPLICANT: Rothschild, Kenneth
; APPLICANT: Gite, Sadanand
; APPLICANT: Olejnik, Jerzy
; TITLE OF INVENTION: Methods
; FILE REFERENCE: AMBER-03951
; CURRENT APPLICATION NUMBER: US/10/264,127
; CURRENT FILING DATE: 2002-10-03
; PRIOR APPLICATION NUMBER: US/09/382,736B
; PRIOR FILING DATE: 1999-08-25
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

US-10-264-127-5
Query Match      66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
Db 1 HHHH 4

US-10-321-204-53
; Sequence 53, Application US/10321204
; Publication No. US20030186871A1
; GENERAL INFORMATION:
; APPLICANT: Waters, Steve
; APPLICANT: Moodie, Shonna
; APPLICANT: Lavan, Brian
; APPLICANT: Gustafson, Thomas A.
; APPLICANT: Metabolex, Inc.
; TITLE OF INVENTION: Compositions and Methods for Diagnosing and Treating
; TITLE OF INVENTION: Diabetes, Insulin Resistance and Dyslipidemia
; FILE REFERENCE: 016325-007310US
; CURRENT APPLICATION NUMBER: US/10/321,204
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: US 60/341,451
; PRIOR FILING DATE: 2001-12-17
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:hexahistidine
; OTHER INFORMATION: affinity tag
US-10-321-204-53

Query Match      66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
Db 1 HHHH 4

US-10-264-127-5
; Sequence 5, Application US/10264127
; Publication No. US20030190643A1
; GENERAL INFORMATION:
; APPLICANT: Rothschild, Kenneth
; APPLICANT: Gite, Sadanand
; APPLICANT: Olejnik, Jerzy
; TITLE OF INVENTION: Methods
; FILE REFERENCE: AMBER-03951
; CURRENT APPLICATION NUMBER: US/10/264,127
; CURRENT FILING DATE: 2002-10-03
; PRIOR APPLICATION NUMBER: US/09/382,736B
; PRIOR FILING DATE: 1999-08-25
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

US-10-264-127-5
Query Match      66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHH 5
Db 1 HHHH 4

US-10-317-428-14
; Sequence 14, Application US/10317428
; Publication No. US20030186414A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Young, N. Martin
; APPLICANT: Waxarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Nucleic Acid That Encodes a Fusion Protein
; FILE REFERENCE: 019633-000811US
; CURRENT APPLICATION NUMBER: US/10/317,428
; CURRENT FILING DATE: 2002-12-11
; PRIOR APPLICATION NUMBER: US 60/069,443
; PRIOR FILING DATE: 1997-12-15
; PRIOR APPLICATION NUMBER: US 09/211,691
; PRIOR FILING DATE: 1998-12-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
```

Query Match 66.7%; Score 4; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 HHHH 5
Db 1 HHHH 4

Search completed: October 21, 2003, 18:49:39
Job time : 42.8667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:19:26 ; Search time 39.7222 Seconds
(without alignments)

32.482 Million cell updates/sec

Title: US-10-057-890A-16

Perfect score: 5

Sequence: 1 GGGGS 5

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1349

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Listing first 150 summaries

Database : SPTREMBL 23:**

1: sp_archaea:**

2: sp_bacteria:**

3: sp_fungi:**

4: sp_human:**

5: sp_invertebrate:**

6: sp_mammal:**

7: sp_mhc:**

8: sp_organelle:**

9: sp_phase:**

10: sp_plant:**

11: sp_rodent:**

12: sp_virus:**

13: sp_vertebrate:**

14: sp_unclassified:**

15: sp_rvirus:**

16: sp_bacteriap:**

17: sp_archaeap:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4	80.0	9	Q9BQ02	Q9BQ02 homo sapien
2	3	60.0	10	Q96QA7	Q96QA7 homo sapien
3	2	40.0	7	Q63480	Q63480 rattus norv
4	2	40.0	7	O55184	O55184 rattus norv
5	2	40.0	8	Q9HDS4	Q9HDS4 aspergillus
6	2	40.0	8	Q15901	Q15901 homo sapien
7	2	40.0	8	Q9UCN4	Q9UCN4 homo sapien
8	2	40.0	8	Q9UMC7	Q9UMC7 homo sapien
9	2	40.0	8	Q8GTG5	Q8GTG5 lycopersico
10	2	40.0	8	Q9QV15	Q9QV15 rattus sp.
11	2	40.0	8	Q60615	Q60615 mus musculus
12	2	40.0	8	Q64971	Q64971 alfalfa mos
13	2	40.0	9	Q99193	Q99193 pseudomonas
14	2	40.0	9	Q9H3Y3	Q9H3Y3 homo sapien
15	2	40.0	9	Q9UQW0	Q9UQW0 homo sapien
16	2	40.0	9	Q8WT19	Q8WT19 drosophila

17	2	40.0	9	6	Q9TRW2	Q9TRW2 oryctolagus
18	2	40.0	9	7	Q31415	Q31415 gallus gall
19	2	40.0	9	8	Q94NB1	Q94NB1 microcebus
20	2	40.0	9	8	Q94NE2	Q94NE2 microcebus
21	2	40.0	9	8	Q94NA9	Q94NA9 daubentonina
22	2	40.0	9	8	Q94XE6	Q94XE6 tectocoris
23	2	40.0	9	8	Q94NB0	Q94NB0 microcebus
24	2	40.0	9	11	Q8RS14	Q8RS14 rattus norv
25	2	40.0	9	11	Q9QWT0	Q9QWT0 mus musculus
26	2	40.0	9	12	Q92766	Q92766 canine dist
27	2	40.0	9	13	Q92009	Q92009 gallus gall
28	2	40.0	10	2	Q9X3M2	Q9X3M2 prochloroco
29	2	40.0	10	2	O52837	O52837 rhizobium 1
30	2	40.0	10	2	Q9R7J8	Q9R7J8 helicobacte
31	2	40.0	10	2	Q51812	Q51812 escherichia
32	2	40.0	10	3	Q9HDS2	Q9HDS2 aspergillus
33	2	40.0	10	3	Q9HDS1	Q9HDS1 aspergillus
34	2	40.0	10	4	Q8NEY9	Q8NEY9 homo sapien
35	2	40.0	10	4	Q9UCU6	Q9UCU6 homo sapien
36	2	40.0	10	6	O9TS42	O9TS42 sus scrofa
37	2	40.0	10	6	O8MP04	O8MP04 ateles belz
38	2	40.0	10	6	O8MJ78	O8MJ78 bos mutus g
39	2	40.0	10	8	Q37103	Q37103 saccharomyc
40	2	40.0	10	8	Q35013	Q35013 meloidogyne
41	2	40.0	10	10	P82938	P82938 hordeum vul
42	2	40.0	10	11	O08622	O08622 rattus norv
43	2	40.0	10	11	O8BHN2	O8BHN2 mus musculus
44	2	40.0	10	13	O8AXT7	O8AXT7 xenopus lae
45	2	40.0	10	15	Q9QKJ0	Q9QKJ0 human immun
46	1	20.0	6	10	P82181	P82181 spinacia ol
47	1	20.0	6	10	P82541	P82541 spinacia ol
48	1	20.0	6	10	P82182	P82182 spinacia ol
49	1	20.0	7	2	O8KMS3	O8KMS3 klebsiella
50	1	20.0	7	2	Q47505	Q47505 escherichia
51	1	20.0	7	2	P70804	P70804 azotobacter
52	1	20.0	7	2	Q47029	Q47029 enterobacte
53	1	20.0	7	2	P72081	P72081 nocardia la
54	1	20.0	7	2	O07354	O07354 synechococc
55	1	20.0	7	8	O8MPY6	O8MPY6 taraxacum (
56	1	20.0	7	8	Q95945	Q95945 saccharomyc
57	1	20.0	7	10	P93233	P93233 lycopersico
58	1	20.0	7	10	O49223	O49223 glycine max
59	1	20.0	7	10	Q9CSB3	Q9CSB3 arabidopsis
60	1	20.0	7	10	P82445	P82445 nicotiana t
61	1	20.0	7	11	Q63668	Q63668 rattus norv
62	1	20.0	7	12	Q65578	Q65578 bovine herp
63	1	20.0	7	13	Q8J320	Q8J320 gallus gall
64	1	20.0	7	13	Q42564	Q42564 fugu rubrip
65	1	20.0	7	15	Q07624	Q07624 rous sarcom
66	1	20.0	8	2	O9AGP4	O9AGP4 athrobacte
67	1	20.0	8	2	Q9R3X0	Q9R3X0 planktothri
68	1	20.0	8	2	Q9R5R2	Q9R5R2 shigella dy
69	1	20.0	8	2	O68485	O68485 klebsiella
70	1	20.0	8	2	Q93SR0	Q93SR0 staphylococ
71	1	20.0	8	2	Q8RJ10	Q8RJ10 streptomyce
72	1	20.0	8	2	O44463	O44463 rhizobiales
73	1	20.0	8	2	P77556	P77556 escherichia
74	1	20.0	8	2	O9ZE29	O9ZE29 buchnera ap
75	1	20.0	8	2	O09258	O09258 synechococc
76	1	20.0	8	2	Q8RSF3	Q8RSF3 lactobacill
77	1	20.0	8	2	Q9X3K1	Q9X3K1 prochloroco
78	1	20.0	8	2	O87471	O87471 haemophilus
79	1	20.0	8	2	O56140	O56140 streptococc
80	1	20.0	8	2	O56246	O56246 thermophili
81	1	20.0	8	2	O56429	O56429 thermus the
82	1	20.0	8	2	Q9R9E0	Q9R9E0 bacillus su
83	1	20.0	8	2	Q35FP2	Q35FP2 pseudomonas
84	1	20.0	8	2	O45889	O45889 clostridias
85	1	20.0	8	2	O9R5R0	O9R5R0 shigella dy
86	1	20.0	8	2	O9R9C2	O9R9C2 borrelia bu
87	1	20.0	8	2	O47273	O47273 escherichia
88	1	20.0	8	2	Q9R4W3	Q9R4W3 enterococcu
89	1	20.0	8	2	Q93454	Q93454 thriobacillu

90 P72221 pseudomonas
 91 O85406 coxiella bu
 92 Q95157 clostridium
 93 P83158 anabaena sp
 94 Q05403 saccharomyc
 95 P87225 saccharomyc
 96 Q9URB9 saccharomyc
 97 P82858 puccinia re
 98 Q15889 homo sapien
 99 Q15894 homo sapien
 100 Q16428 homo sapien
 101 Q15895 homo sapien
 102 Q9Y4X6 homo sapien
 103 Q15890 homo sapien
 104 Q15898 homo sapien
 105 Q15893 homo sapien
 106 Q15888 homo sapien
 107 Q16468 homo sapien
 108 Q9HCQ0 homo sapien
 109 Q9P285 homo sapien
 110 Q8TF70 homo sapien
 111 Q9P0K3 homo sapien
 112 Q60773 homo sapien
 113 Q81UB8 homo sapien
 114 Q9TWH6 perinereis
 115 Q94623 manduca sex
 116 Q02032 lytechinus
 117 Q94695 physarum po
 118 P82685 periplaneta
 119 P82686 periplaneta
 120 P82687 periplaneta
 121 P82688 periplaneta
 122 P82689 periplaneta
 123 Q9TT78 canis famil
 124 Q9BFA7 macroscelid
 125 Q18854 canis famil
 126 Q9XSY1 canis famil
 127 Q9MYL5 pongo pygma
 128 Q9GMB3 lagenorhync
 129 Q28866 megaptera n
 130 Q9TRY3 sus sp. ins
 131 Q29801 homo sapien
 132 Q29812 homo sapien
 133 Q29810 homo sapien
 134 Q95213 oryctolagus
 135 Q29797 homo sapien
 136 Q19957 gossypium h
 137 Q9MSX1 jurinea hum
 138 Q93963 psathyrosta
 139 Q94VC1 varanus rud
 140 Q36898 nicotiana p
 141 P92384 hordeum mur
 142 P93973 eremopyrum
 143 P92215 amblyopyrum
 144 Q9TZY3 begonia for
 145 P93957 festucopsis
 146 P92222 bromus iner
 147 P92388 henrardia p
 148 P92441 thinopyrum
 149 Q9GD00 mascula mad
 150 P92404 lophopyrum

ALIGNMENTS

RESULT 1
 98BQ02
 D Q9BQ02 PRELIMINARY; PRT; 9 AA.
 C Q9BQ02;
 T 01-JUN-2001 (TrEMBLrel. 17, Created)
 T 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 T 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)

DE BA28009.2.1 (Novel protein (Isoform 1)) (Fragment) .
 GN DJ63IM13.5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lloyd D.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EXBL: AL121582; CAC36068.1; .
 FT NON_TER 1 1
 FT NON_TER 9 9
 SQ SEQUENCE 9 AA; 745 MW; C18E7878772735BD CRC64;
 Query Match 80.0%; Score 4; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GGGG 4
 Db |||||
 6 GGGG 9
 RESULT 2
 98BQ07
 ID Q96QA7 PRELIMINARY; PRT; 10 AA.
 AC Q96QA7;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE BA151A2.1 (Cdc42 guanine exchange factor (GEF) 9 (Collybistin, PEM-2,
 HPEM-2, KIAA0424)) (Fragment).
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Whitehead S.;
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL451106; CAC88408.1; .
 FT NON_TER 10 10
 SQ SEQUENCE 10 AA; 1122 MW; 39925CB878640043 CRC64;
 Query Match 60.0%; Score 3; DB 4; Length 10;
 Best Local Similarity 100.0%; Pred. No. 4.2e+03;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 3 GGS 5
 Db |||||
 6 GGS 8
 RESULT 3
 98BQ480
 ID Q63480 PRELIMINARY; PRT; 7 AA.
 AC Q63480;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE TR4-NS orphan receptor (fragment).
 GN TR4.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=96198747; PubMed=8612486;
 RA Yoshikawa T.; Makino S.; Gao X.M.; Xing G.Q.; Chuang D.M.;
 RA Detera-Wadleigh S.D.;

RT "Splice variants of rat TR4 orphan receptor: differential expression
 RT of novel sequences in the 5'-untranslated region and C-terminal
 RL domain.";
 RL Endocrinology 137:1562-1571(1996).
 DR EMBL; U59125; AAB02827.1; -.
 KW Receptor.
 FT NON_TER 1
 SQ SEQUENCE 7 AA; 758 MW; 672AA87864005350 CRC64;

Query Match 40.0%; Score 2; DB 11; Length 7;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
 Db 4 GG 5

RESULT 4

OS5184 PRELIMINARY; PRT; 7 AA.
 ID O55184
 AC O55184
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Orphan receptor TR4-NS (Fragment).
 GN TR4.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RX MEDLINE=96198747; PubMed=8612486;
 RA Yoshikawa T., Makino S., Gao X.M., Xing G.Q., Chuang D.M.,
 RA Detera-Wadleigh S.D.;
 "Splice variants of rat TR4 orphan receptor: differential expression
 RT of novel sequences in the 5'-untranslated region and C-terminal
 RT domain.";
 RL Endocrinology 137:1562-1571(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RX MEDLINE=96299786; PubMed=8661150;
 RA Yoshikawa T., DuPont B.R., Leach R.J., Detera-Wadleigh S.D.;
 RT "New variants of the human and rat nuclear hormone receptor; TR4:
 RT expression and chromosomal localization of the human gene.";
 RL Genomics 35:361-366(1996).
 DR EMBL; U59454; AAB91433.1; -.
 KW Receptor.
 FT NON_TER 1
 SQ SEQUENCE 7 AA; 663 MW; 6DDAA878EB05350 CRC64;

Query Match 40.0%; Score 2; DB 11; Length 7;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
 Db 4 GG 5

RESULT 5

Q9HDS4 PRELIMINARY; PRT; 8 AA.
 ID Q9HDS4
 AC Q9HDS4
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DE TrpC polypeptide (Fragment).
 GN TrpC.
 OS Aspergillus flavus.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5059;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A55;
 RA Geiser D.M., Dörner J.W., Horn B.W., Taylor J.W.;
 RT "The phylogenetics of mycotoxin and sclerotium production in
 RT Aspergillus flavus and Aspergillus oryzae.";
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF261861; AAG16135.1; -.
 KW Polyprotein.
 FT NON_TER 8
 SQ SEQUENCE 8 AA; 807 MW; F3B2C72AB5B87DD6 CRC64;

Query Match 40.0%; Score 2; DB 3; Length 8;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
 Db 3 GS 4

RESULT 6

Q15901 PRELIMINARY; PRT; 8 AA.
 ID Q15901
 AC Q15901
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE (Clone XP7B11B) (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RA Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
 RA Coolbaugh M.I., Chinault C.A., Baldini A., Lindsey E.A., Zhao Z.-Y.,
 RA Caskey C.T.H.;
 RT "Isolation of chromosome-specific genes by reciprocal probing of
 RT arrayed cDNAs and cosmid libraries.";
 RL Hum. Mol. Genet. 0:0-0(1995).
 DR EMBL; L32080; AAA73891.1; -.
 FT NON_TER 1
 FT NON_TER 8
 SQ SEQUENCE 8 AA; 860 MW; 37D72878676729CB CRC64;

Query Match 40.0%; Score 2; DB 4; Length 8;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
 Db 5 GG 6

RESULT 7

Q9UCN4 PRELIMINARY; PRT; 8 AA.
 ID Q9UCN4
 AC Q9UCN4
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE Cell-surface heparin/HEPARANSULFATE-binding protein peptide 3
 DE (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]


```

RP SEQUENCE.
RX MEDLINE=92291065; PubMed=1601862;
RA Raboudi N., Julian J., Rohde L.H., Carson D.D.;
RT "Identification of cell-surface heparin/heparan sulfate-binding
RT proteins of a human uterine epithelial cell line (RL95).";
RL J. Biol. Chem. 267:11930-11939(1992).
SQ SEQUENCE 8 AA; 689 MW; 8088733DD33DD87D CRC64;

Query Match 40.0%; Score 2; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
DB 7 GS 8

RESULT 8
Q9UMC7 PRELIMINARY; PRT; 8 AA.
AC Q9UMC7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE SHMT protein (Fragment).
DE SHMT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Chave K.J., Snell K., Sanders P.G.;
RT "Isolation and characterisation of human genomic sequences encoding
RT cytosolic serine hydroxymethyltransferase.";
RL Biochem. Soc. Trans. 25:53-53(1997).
DR EMBL; Y14492; CAB54844.1; -.
FT NON_TER 1 1
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 868 MW; 7C205721E44AB588 CRC64;

Query Match 40.0%; Score 2; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
DB 1 GS 2

RESULT 9
Q8GTG5 PRELIMINARY; PRT; 8 AA.
AC Q8GTG5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE WRKY transcription factor Iid-1 splice variant 2 (Fragment).
DE Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; Lamiales; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Money Maker; TISSUE=Leaf;
RA Cagna G., Boes A., Ulker B., Zhou A., Wanke D., Somsich I.E.;
RT "Comparison of WRKY group II transcription factors from plants.";
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY157059; AAN71729.1; -.
FT NON_TER 1 1
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 917 MW; 59177B4775B87330 CRC64;

```

```

Query Match 40.0%; Score 2; DB 10; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
DB 3 GS 4

RESULT 10
Q9QVI5 PRELIMINARY; PRT; 8 AA.
AC Q9QVI5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 18 kDa cell growth factor (Fragment).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE.
RX MEDLINE=92028975; PubMed=1656977;
RA Milner P.G.;
RT "Simian sarcoma virus transformation of normal rat kidney fibroblasts
RT is associated with markedly increased basic fibroblast growth factor
RT expression.";
RL Biochem. Biophys. Res. Commun. 180:423-430(1991).
FT NON_TER 1 1
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 807 MW; B0787AAB07673AFA CRC64;

Query Match 40.0%; Score 2; DB 11; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GS 2
DB 7 GS 8

RESULT 11
Q60615 PRELIMINARY; PRT; 8 AA.
AC Q60615;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE Apolipoprotein A-II (Fragment).
GN APOA2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=94319082; PubMed=8043949;
RA Ko M.S., Wang X., Horton J.H., Hagen M.D., Takahashi N., Maezaki Y.,
RA Nadeau J.H.;
RT "Genetic mapping of 40 cDNA clones on the mouse genome by PCR.";
RL Mamm. Genome 5:349-355(1994).
DR EMBL; U05691; AAB60462.1; -.
DR MGD; MGI:88050; APOA2.
KW Lipoprotein.
FT NON_TER 1 1
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 720 MW; 783DDDC5B861AB18 CRC64;

Query Match 40.0%; Score 2; DB 11; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 4 GS 5
 Db 4 GS 5

RESULT 12
 Q64971 PRELIMINARY; PRT; 8 AA.
 AC Q64971;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Putative ORF (Fragment).
 DE OS Alfalfa mosaic virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
 OC Alfamovirus.
 OX NCBI_TaxID=12321;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=81124289; PubMed=6927843;
 RA Koper-Zwarthoff E.C., Brederode F.T.M., Veeneman G., van Boom J.H.,
 RA Bol J.F.;
 RT "Nucleotide sequences at the 5'-termini of the alfalfa mosaic virus
 RT RNAs and the interstronic junction in RNA 3.";
 RL Nucleic Acids Res. 8:5635-5647(1980).
 DR EMBL; V00047; CAA23416.1; -.
 FT NON TER 1
 SQ SEQUENCE 8 AA; 917 MW; 69D40B0775A3365B8 CRC64;

Query Match 40.0%; Score 2; DB 12; Length 8;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
 Db 1 GS 2

RESULT 13

Q99193 PRELIMINARY; PRT; 9 AA.
 AC Q99193;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE RpoB beta-subunit of RNA polymerase (Fragment).
 OS Pseudomonas putida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Borodin A.M., Danilovich A.V., Allikmets R.L., Rostapshov V.M.,
 RA Chernov I.P., Azhikina T.L., Monastyrskaya S., Sverdlov D.;
 RT "Nucleotide sequence of the rpoB gene coding for the beta-subunit of
 RT RNA polymerase in Pseudomonas putida.";
 RL Dokl. Biochem. 302:1261-1265(1988).
 DR EMBL; X15849; CAA33847.1; -.
 FT NON TER 9
 SQ SEQUENCE 9 AA; 852 MW; 5B416776DC76727 CRC64;

Query Match 40.0%; Score 2; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GS 2
 Db 8 GS 9

RESULT 14

Q9H3Y3 PRELIMINARY; PRT; 9 AA.
 ID Q9H3Y3

AC Q9H3Y3;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE DJ839B11.1 (Novel protein with a kunitz/bovine pancreatic trypsin
 DE inhibitor domain and WAP-type (Whey acidic protein) 'four-disulfide
 DE core' domain) (Fragment).
 GN DJ461P17.1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lloyd D.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL121778; CAB76844.1; -.
 FT NON TER 9
 SQ SEQUENCE 9 AA; 1036 MW; 2C417B01B412D1B3 CRC64;

Query Match 40.0%; Score 2; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GS 2
 Db 8 GS 9

RESULT 15

Q9UQW0 PRELIMINARY; PRT; 9 AA.
 AC Q9UQW0;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE Prolactin precursor (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84182507; PubMed=6325171;
 RA Truong A.T., Duez C., Belayew A., Renard A., Pictet R., Bell G.I.,
 RA Martial J.A.;
 RT "Isolation and characterization of the human prolactin gene.";
 RL EMBO J. 3:429-437(1984).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93076813; PubMed=1332868;
 RA Peers B., Naida A.M., Monget P., Voz M.L., Belayew A., Martial J.A.;
 RT "Binding of a 100-kDa ubiquitous factor to the human prolactin
 RT promoter is required for its basal and hormone-regulated activity.";
 RL Eur. J. Biochem. 210:53-58(1992).
 DR EMBL; X00368; CAA25108.1; -.
 KW Signal.
 FT SIGNAL 1 8 POTENTIAL.
 FT NON TER 9
 SQ SEQUENCE 9 AA; 1060 MW; 0A1A6775B8733054 CRC64;

Query Match 40.0%; Score 2; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
 Db 5 GS 6

RESULT 16

Q8WTI9 PRELIMINARY; PRT; 9 AA.
 ID Q8WTI9

```

AC Q8WTI9;
DE 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Mod(mdg4) protein (Fragment).
DE MOD(MDG4) OR CG7836 OR CG8076 OR CG15500 OR CG15802 OR
GN CG18151 OR CG32491.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=mod(mdg4)ui; TRANSPOSON-Stalker;
RX MEDLINE=21490778; PubMed=11604507;
RA Wei W., Brennan M.D.;
RT "The gypsy insulator Can Act as a Promoter-Specific Transcriptional
RT Stimulator.";
RL Mol. Cell. Biol. 21:7714-7720(2001).
DR EMBL: AF214650; AAL33875.1; -.
DR FlyBase; FBgn0002781; mod(mdg4).
FT NON_TER 1
SQ SEQUENCE 9 AA; 946 MW; 3E5F49C0572322C8 CRC64;

Query Match 40.0%; Score 2; DB 5; Length 9;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
DB ||
1 GG 2

RESULT 17
Q9TRW2 PRELIMINARY; PRT; 9 AA.
ID Q9TRW2;
AC Q9TRW2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CALDESMON=PHOSPHORYLATION site (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OC NCBI_TaxID=9986;
RN [1]
RP SEQUENCE.
RX MEDLINE=91378498; PubMed=1898046;
RA Ikebe M., Hornick T.;
RT "Determination of the phosphorylation sites of smooth muscle caldesmon
RT by protein kinase C.";
RL Arch. Biochem. Biophys. 288:538-542(1991).
FT NON_TER 1
SQ SEQUENCE 9 AA; 1018 MW; 8C901B10533735A5 CRC64;

Query Match 40.0%; Score 2; DB 6; Length 9;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
DB ||
1 GS 2

RESULT 18
Q31415 PRELIMINARY; PRT; 9 AA.
ID Q31415;
AC Q31415;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-JAN-1999 (TrEMBLrel. 09, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

```

```

DE MHC class I antigen (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Kanki T., Kuwasawa N., Sekiya Y., Ichikawa Y.;
RT "Responsive expression of a MHC class I epitope and genes following
RT Marek's disease virus infection.";
RL Submitted (JAN-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL: D90399; BAA14395.1; -.
FT NON_TER 1
SQ SEQUENCE 9 AA; 859 MW; 8A55A76455861B5 CRC64;

Query Match 40.0%; Score 2; DB 7; Length 9;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
DB ||
4 GS 5

RESULT 19
Q94NB1 PRELIMINARY; PRT; 9 AA.
ID Q94NB1;
AC Q94NB1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Cytochrome oxidase subunit III (Fragment).
GN COIII.
OS Microcebus ravelobensis.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Strepsirhini; Cheirogaleidae;
OC Microcebus.
OX NCBI_TaxID=122231;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JP299, and JP301;
RX MEDLINE=21184272; PubMed=11286490;
RA Pastorini J., Martin R.D., Ehresmann P., Zimmermann E., Forstner M.R.;
RT "Molecular phylogeny of the lemur family Cheirogaleidae (primates)
RT based on mitochondrial DNA sequences.";
RL Mol. Phylogenet. Evol. 19:45-56(2001).
DR EMBL: AF224630; AAK70571.1; -.
DR EMBL: AF224631; AAK70575.1; -.
KW Mitochondrion.
FT NON_TER 1
SQ SEQUENCE 9 AA; 1160 MW; D5C563636B5045A2 CRC64;

Query Match 40.0%; Score 2; DB 8; Length 9;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
DB ||
8 GS 9

RESULT 20
Q94NB2 PRELIMINARY; PRT; 9 AA.
ID Q94NB2;
AC Q94NB2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Cytochrome oxidase subunit III (fragment).
GN COIII.
OS Microcebus murinus (Lesser mouse lemur).

```

```

OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Strepsirhini; Cheirogaleidae;
OC Microcebus.
OX NCBI_TaxID=30608;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JP285, JP288, JP289, JP292, JP308, and JP313;
RX MEDLINE=21184272; PubMed=11286490;
RA Pastorini J., Martin R.D., Ehresmann P., Zimmermann E., Forstner M.R.;
RT "Molecular phylogeny of the lemur family Cheirogaleidae (primates)
RL Mol. Phylogenet. Evol. 19:45-56(2001).
DR EMBL; AF224624; AAK70547.1; -.
DR EMBL; AF224625; AAK70551.1; -.
DR EMBL; AF224626; AAK70555.1; -.
DR EMBL; AF224627; AAK70559.1; -.
DR EMBL; AF224628; AAK70563.1; -.
DR EMBL; AF224629; AAK70567.1; -.
KW Mitochondrion.
FT NON_TER 1 1
SQ SEQUENCE 9 AA; 1160 MW; D5C563636B5045A2 CRC64;

Query Match 40.0%; Score 2; DB 8; Length 9;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
DB 8 GS 9

RESULT 21
Q94NA9 PRELIMINARY; PRT; 9 AA.
AC Q94NA9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Cytochrome oxidase subunit III (Fragment).
GN COIII.
OS Daubentonia madagascariensis (Aye-aye).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Strepsirhini; Daubentonidae;
OC Daubentonia.
OX NCBI_TaxID=31869;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JP119, and JP120;
RX MEDLINE=21184272; PubMed=11286490;
RA Pastorini J., Martin R.D., Ehresmann P., Zimmermann E., Forstner M.R.;
RT "Molecular phylogeny of the lemur family Cheirogaleidae (primates)
RL Mol. Phylogenet. Evol. 19:45-56(2001).
DR EMBL; AF224641; AAK70615.1; -.
DR EMBL; AF224642; AAK70619.1; -.
KW Mitochondrion.
FT NON_TER 1 1
SQ SEQUENCE 9 AA; 1160 MW; D5C563636B5045A2 CRC64;

Query Match 40.0%; Score 2; DB 8; Length 9;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
DB 8 GS 9

RESULT 22
Q94XE6 PRELIMINARY; PRT; 9 AA.
ID Q94XE6

```

```

AC Q94XE6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Cytochrome c oxidase subunit III (Fragment).
GN COX3.
OS Tectocoris diopthalmus (cotton harlequin bug).
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Paraneoptera; Hemiptera; Euhemiptera; Heteroptera;
OC Panheteroptera; Pentatomomorpha; Pentatomidae; Pentatomidae;
OC Tectocoris.
OX NCBI_TaxID=159956;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21396409; PubMed=11504862;
RA Shao R., Campbell N.J.H., Schmidt E.R., Barker S.C.;
RT "Increased rate of gene rearrangement in the mitochondrial genomes of
RL three orders of hemipteroid insects."
RL Mol. Biol. Evol. 18:1828-1832(2001).
DR EMBL; AF335990; AAK55283.1; -.
KW Mitochondrion.
FT NON_TER 1 1
SQ SEQUENCE 9 AA; 1206 MW; A2C563636B5041A6 CRC64;

Query Match 40.0%; Score 2; DB 8; Length 9;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
DB 8 GS 9

RESULT 23
Q94NB0 PRELIMINARY; PRT; 9 AA.
ID Q94NB0
AC Q94NB0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Cytochrome oxidase subunit III (Fragment).
GN COIII.
OS Microcebus rufus (brown mouse lemur).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Strepsirhini; Cheirogaleidae;
OC Microcebus.
OX NCBI_TaxID=122332;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JP309, JP315, JP316, and JP317;
RX MEDLINE=21184272; PubMed=11286490;
RA Pastorini J., Martin R.D., Ehresmann P., Zimmermann E., Forstner M.R.;
RT "Molecular phylogeny of the lemur family Cheirogaleidae (primates)
RL Mol. Phylogenet. Evol. 19:45-56(2001).
DR EMBL; AF224636; AAK70595.1; -.
DR EMBL; AF224637; AAK70599.1; -.
DR EMBL; AF224638; AAK70603.1; -.
DR EMBL; AF224639; AAK70607.1; -.
KW Mitochondrion.
FT NON_TER 1 1
SQ SEQUENCE 9 AA; 1160 MW; D5C563636B5045A2 CRC64;

Query Match 40.0%; Score 2; DB 8; Length 9;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
DB 8 GS 9

```

```

DE F. Fusion protein (Fragment).
GN
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Dog #5526/89;
RA Liermann H., Harder T., Haas L.;
RT "Genetic analysis of the central untranslated genome region and the
RT proximal coding part of the F gene of wild-type and vaccine distemper
RT morbilliviruses".
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF026237; AAC09167.1; -.
FT NON_TER 9
SQ SEQUENCE 9 AA; 1011 MW; F281732760533441 CRC64;

Query Match 40.0%; Score 2; DB 12; Length 9;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
DB ||
8 GS 9

RESULT 27
Q92309
ID Q92309 PRELIMINARY; PRT; 9 AA.
AC Q92309;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE C-REL protein (fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=9113738; PubMed=2284104;
RA Hannink M., Temin H.M.;
RT "Structure and autoregulation of the c-rel promoter.";
RL Oncogene 5:1843-1850(1990).
RN [2]
RP SEQUENCE FROM N.A.
RA Hannink M., Temin H.M.;
RL Oncogene 0:0-0(1990).
DR EMBL; X56440; CAA39822.1; -.
DR EMBL; X56515; CAA39866.1; -.
FT NON_TER 9
SQ SEQUENCE 9 AA; 805 MW; DE317DD87865A2CD CRC64;

Query Match 40.0%; Score 2; DB 13; Length 9;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
DB ||
5 GG 6

RESULT 28
Q9X3M2
ID Q9X3M2 PRELIMINARY; PRT; 10 AA.
AC Q9X3M2;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE Cytochrome b (fragment).
GN PEB.

DE F. Fusion protein (Fragment).
GN
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Dog #5526/89;
RA Liermann H., Harder T., Haas L.;
RT "Genetic analysis of the central untranslated genome region and the
RT proximal coding part of the F gene of wild-type and vaccine distemper
RT morbilliviruses".
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF026237; AAC09167.1; -.
FT NON_TER 9
SQ SEQUENCE 9 AA; 1011 MW; F281732760533441 CRC64;

Query Match 40.0%; Score 2; DB 11; Length 9;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
DB ||
6 GG 7

RESULT 25
Q9QWTO
ID Q9QWTO PRELIMINARY; PRT; 9 AA.
AC Q9QWTO;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Proteinase 3 (fragment).
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sv129 D3;
RA Lutz P.G., Moog-Lutz C., Houzel-Charavel A., Cayre Y.E.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ007030; CAA07429.1; -.
DR MCD; MGI:893580; Ptn3.
FT NON_TER 9
SQ SEQUENCE 9 AA; 937 MW; C91875A77B45B87D CRC64;

Query Match 40.0%; Score 2; DB 11; Length 9;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
DB ||
3 GS 4

RESULT 26
Q92766
ID Q92766 PRELIMINARY; PRT; 9 AA.
AC Q92766;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)

```

```

OS Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1220;
RN [1]
RP SEQUENCE FROM N.A.
RA Urbach E., Chisholm S.W.;
RT "Genetic diversity in Prochlorococcus populations flow cytometrically
RT sorted from the Sargasso Sea and Gulf Stream.";
RL Limnol. Oceanogr. 43:1615-1630(1998).
DR EMBL; AF070219; AAD23249.1; -.
FT NON_TER 1
SQ SEQUENCE 10 AA; 1076 MW; 75CA5CB05866D324 CRC64;

Query Match 40.0%; Score 2; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GS 5
Db 9 GS 10

RESULT 29
QS2837 PRELIMINARY; PRT; 10 AA.
AC QS2837;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE FixX protein (10 AA) (Fragment).
OS Rhizobium leguminosarum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=384;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PRE;
RX MEDLINE=90136072; PubMed=2693897;
RA Roelvink P.W., Montelez J.G.J., van Kammen A., van den Bos R.C.;
RT "Nucleotide sequence of the regulatory nifA gene of Rhizobium
RT leguminosarum PRE: transcriptional control sites and expression in
RT Escherichia coli.";
RL Mol. Microbiol. 3:1441-1447(1989).
DR EMBL; X17073; CAA34923.1; -.
FT NON_TER 1
SQ SEQUENCE 10 AA; 1037 MW; 24A5593732C879C8 CRC64;

Query Match 40.0%; Score 2; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GG 2
Db 2 GG 3

RESULT 30
Q9R7J8 PRELIMINARY; PRT; 10 AA.
AC Q9R7J8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE Vacuolating cytotoxin (Fragment).
OS VACA.
GN Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kobe 500;

```

```

RA Shitasaka D.;
RT "Helicobacter pylori vacA gene, strain Kobe 500, partial cds.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB017599; BAAJ3412.1; -.
FT NON_TER 1
FT NON_TER 10
SQ SEQUENCE 10 AA; 1018 MW; 414390C76879CDD7 CRC64;

Query Match 40.0%; Score 2; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GS 5
Db 9 GS 10

RESULT 31
QS1812 PRELIMINARY; PRT; 10 AA.
AC QS1812;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE DNA helicase (fragment).
GN TRAI.
OS Escherichia coli.
OG Plasmid F.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K-12;
RX MEDLINE=92009201; PubMed=1916281;
RA Cram D.S., Loh S.M., Cheah K.C., Skurray R.A.;
RT "Sequence and conservation of genes at the distal end of the transfer
RT region on plasmids F and R6-5.";
RL Gene 104:85-90(1991).
DR EMBL; M38047; AAA98090.1; -.
KW Helicase; Plasmid.
FT NON_TER 1
SQ SEQUENCE 10 AA; 1088 MW; 59A3047731A33B13 CRC64;

Query Match 40.0%; Score 2; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GG 2
Db 8 GG 9

RESULT 32
Q9HDS2 PRELIMINARY; PRT; 10 AA.
AC Q9HDS2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE TrpC polyprotein (fragment).
GN TRPC.
OS Aspergillus flavus.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5059;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TX19_21S;
RA Geiser D.M., Dörner J.W., Horn B.W., Taylor J.W.;
RT "The phylogenetics of mycotoxin and sclerotium production in
RT Aspergillus flavus and Aspergillus oryzae.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

```

DR EMBL; AF261865; AAG16139.1; --

KW Polyprotein.

FT NON_TER 1 1

SQ SEQUENCE 10 AA; 997 MW; 8FBEB4EAB2C72AB5 CRC64;

Query Match 40.0%; Score 2; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5

DB 2 GS 3

RESULT 33

Q9HDS1

ID Q9HDS1

AC Q9HDS1

DT 01-MAR-2001

DT 01-MAR-2001

DE TrpC polyprotein (Fragment).

OS TrpC.

OC Aspergillus flavus.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

OC NCBI_TaxID=5059;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=F35;

RA Geiser D.M., Dörner J.W., Horn B.W., Taylor J.M.;

RT "The phylogenetics of mycotoxin and sclerotium production in

RT Aspergillus flavus and Aspergillus oryzae."

RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF261870; AAG16144.1; --

KW Polyprotein.

FT NON_TER 1 1

SQ SEQUENCE 10 AA; 1034 MW; 8FB1E4EB02C72AB5 CRC64;

Query Match 40.0%; Score 2; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5

DB 2 GS 3

RESULT 34

Q8NEY9

ID Q8NEY9

AC Q8NEY9

DT 01-OCT-2002

DT 01-OCT-2002

DE Type II hair-specific keratin (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OC NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Bairwa N.K., Bamezai R.;

RT "Study of polymorphism in the regulatory region of KRTHB6 in

RT monilethrix samples."

RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY037552; AAK68688.1; --

KW Keratin.

FT NON_TER 10 10

SQ SEQUENCE 10 AA; 1034 MW; 9B53417EAB45B87E CRC64;

Query Match 40.0%; Score 2; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2

DB 8 GG 9

RESULT 35

Q9UCU6

ID Q9UCU6

AC Q9UCU6

DT 01-MAY-2000

DT 01-MAY-2000

DE Cathepsin G-PLASMA-membrane associated protease (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OC NCBI_TaxID=9606;

RN [1]

RP SEQUENCE.

RX MEDLINE=91318179; PubMed=1861080;

RA Maison C.M., Villiers C.L., Colomb M.G.;

RT "Proteolysis of C3 on U937 cell plasma membranes. Purification of

RT cathepsin G."

RL J. Immunol. 147:921-926(1991).

FT NON_TER 1 1

SQ SEQUENCE 10 AA; 1121 MW; 20805915BB040878 CRC64;

Query Match 40.0%; Score 2; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2

DB 3 GG 4

RESULT 36

Q9TS42

ID Q9TS42

AC Q9TS42

DT 01-MAY-2000

DT 01-MAY-2000

DE OESTRADIOL-RECEPTOR=P2 peptide (Fragment).

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OC NCBI_TaxID=9823;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=91291128; PubMed=2064608;

RA Thole H.H., Jungblut P.W., Jakob F.;

RT "The proton-driven dissociation of oestradiol-receptor dimers as a

RT preparative tool. Isolation of a 32 kDa fragment from porcine uteri

RT and assignment of C-terminal origin by partial sequencing."

RL Biochem. J. 276:709-714(1991).

FT NON_TER 1 1

SQ SEQUENCE 10 AA; 992 MW; D95E047B1451B76D CRC64;

Query Match 40.0%; Score 2; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2

DB 8 GG 9

```

RESULT 37
Q8WP04 PRELIMINARY; PRT; 10 AA.
AC Q8WP04;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Oculocutaneous albinism type II (Fragment).
GN OCA2.
OS Ateles belzebuth chamek (Chamek spider monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Ateleinae; Ateles.
OX NCBI_TaxID=118643;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21636947; PubMed=11778686;
RA Seanez H.N., Lima C.R., Lemos B., Bonvicino C.R., Moreira M.A.M.,
RA Canavez F.C.;
RT "Gene assignment in Ateles paniscus chamek (Platyrrhini, Primates).
RT Allocation of 18 markers of human syntenic groups 1,2,7,14,15,17 and
RT 22."
RL Chromosome Res. 9:631-639(2001).
EL EMBL; AF375652; AAL31489.1; -.
DR EMBL; AF375652; AAL31489.1; -.
FT NON TER 1
SQ SEQUENCE 10 AA; 901 MW; 22DF477DD87EA588 CRC64;

Query Match 40.0%; Score 2; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
DB 3 GS 4

RESULT 38
Q8MJ78 PRELIMINARY; PRT; 10 AA.
AC Q8MJ78;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Prolactin (Fragment).
GN PRL.
OS Bos mutus grunniens (Yak).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=30521;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Blood;
RA Ou J.T., Jr., Zhong J.C. Sr., Chen Z.H. III, Zhao S.J. IV, Bai W.L. V.;
RT "Cloning, sequencing and polymorphism analysis on prolactin gene of
RT yak."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AFS16697; AAM61895.1; -.
FT NON TER 10
SQ SEQUENCE 10 AA; 1024 MW; 93781FC5A5B87325 CRC64;

Query Match 40.0%; Score 2; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
DB 5 GS 6

RESULT 39
Q37103 PRELIMINARY; PRT; 10 AA.
AC Q37103;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 1.1 kDa protein.
OS Saccharomyces cerevisiae (Baker's yeast).
OG Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=WILD TYPE STRAIN A;
RX MEDLINE=85155510; PubMed=6397406;
RA de Zamaroczy M., Faugeron-Fonty G., Baldacci G., Courso R.,
RA Bernardi G.;
RT "The ori sequences of the mitochondrial genome of a wild-type yeast
RT strain: number, location, orientation and structure."
RL Gene 32:439-457(1984).
DR EMBL; K03305; AAA69858.1; -.
DR EMBL; K03296; AAA69852.1; -.
DR EMBL; K03297; AAA69853.1; -.
DR EMBL; K03300; AAA69854.1; -.
KW Hypothetical protein; Mitochondrion.
SQ SEQUENCE 10 AA; 1141 MW; 1FA38165B6933334 CRC64;

Query Match 40.0%; Score 2; DB 8; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
DB 8 GS 9

RESULT 40
Q35013 PRELIMINARY; PRT; 10 AA.
AC Q35013;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Mitochondrial transfer RNA His, 16S ribosomal RNA (16S rRNA) genes,
DE ND3 (Fragment).
OS Meloidogyne javanica (Root-knot nematode).
OG Mitochondrion.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
OC Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.
OX NCBI_TaxID=6303;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97153821; PubMed=9000752;
RA Hugall A., Stanton J., Moritz C.;
RT "Evolution of the AT-rich mitochondrial DNA of the root knot nematode,
RT Meloidogyne hapla."
RL Mol. Biol. Evol. 14:40-48(1997).
DR EMBL; L76261; AAA98776.1; -.
KW Mitochondrion.
FT NON TER 1
SQ SEQUENCE 10 AA; 1080 MW; 556BBD35B8769697 CRC64;

Query Match 40.0%; Score 2; DB 8; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
DB 6 GS 7

RESULT 41
P82938

```


ID P82938 PRELIMINARY; PRT; 10 AA.
AC P82938;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Unknown endosperm protein C (Fragment).
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Hordeum.
NCBI_TaxID=4513;
[1] _
SEQUENCE.
STRAIN=cv. BOMI; TISSUE=Starchy endosperm;
MEDLINE=21088911; PubMed=11271488;
KRISTOFERSEN H.E., FLINGSRUD R.;
"Separation and characterization of basic barley seed proteins.";
Electrophoresis 21:3693-3700(2000).
-1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
PROTEIN IS: 8.5-9.0, ITS MW IS: 11.9 KDA.
NON_TER 10 10
SEQUENCE 10 AA; 1053 MW; 92562DC40AA87AAE CRC64;
Query Match 40.0%; Score 2; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GG 2
Db 1 GG 2
RESULT 42
ID Q08622 PRELIMINARY; PRT; 10 AA.
AC Q08622;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 1.1 kDa protein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
[1] _
SEQUENCE FROM N.A.
MEDLINE=97322348; PubMed=9177193;
PULS AW., SCHMIDT S., GRAWE F., STABEL S.;
"Interaction of protein kinase C zeta with ZIP, a novel protein kinase
C-binding protein.";
Proc. Natl. Acad. Sci. U.S.A. 94:6191-6196(1997).
DR EMBL; Y08355; CAA69641.1; -
KW Hypothetical protein.
SEQUENCE 10 AA; 1050 MW; 3D703E272EA5B728 CRC64;
Query Match 40.0%; Score 2; DB 11; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GG 2
Db 2 GG 3
RESULT 43
ID Q8BHN2 PRELIMINARY; PRT; 10 AA.
AC Q8BHN2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Methylentetrahydrofolate reductase short isoform (Fragment).
GN MTHFR.

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1] _
SEQUENCE FROM N.A.
RC STRAIN=129/SV, and BALB/c;
RX MEDLINE=22257759; PubMed=12370778;
TRAN P., LECLERC D., CHAN M., PAI A., HIOU-TIM P., WU Q., GOYETTE P.,
ARTIGAS C., MILOS R., ROZEN R.;
"Multiple transcription start sites and alternative splicing in the
methylentetrahydrofolate reductase gene result in two enzyme
isoforms.";
Mamm. Genome 13:483-492(2002).
DR EMBL; AY046557; AAL17641.1; -
DR EMBL; AF398931; AAN40867.1; -
FT NON_TER 10 10
SEQUENCE 10 AA; 1007 MW; 01695CB8640DB14 CRC64;
Query Match 40.0%; Score 2; DB 11; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 GS 5
Db 7 GS 8
RESULT 44
ID Q8AXT7 PRELIMINARY; PRT; 10 AA.
AC Q8AXT7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Homeo box-containing protein (Fragment).
GN XANFL.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8355;
RN [1] _
SEQUENCE FROM N.A.
RX MEDLINE=94220989; PubMed=7909469;
RA SAMOKHVALOV I.M., SEMENOVA N.A., NIKOLAEV A.I., BELIAVSKII A.V.;
"Genomic structure of the homeo box-containing gene XANFL.";
Dokl. Akad. Nauk SSSR 334:522-525(1994).
DR EMBL; S71802; AAD14094.2; -
FT NON_TER 10 10
SEQUENCE 10 AA; 991 MW; 3325817326C2CDC7 CRC64;
Query Match 40.0%; Score 2; DB 13; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 GS 5
Db 8 GS 9
RESULT 45
ID Q9QKJ0 PRELIMINARY; PRT; 10 AA.
AC Q9QKJ0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
NCBI_TaxID=11676;

```

RN  SEQUENCE FROM N.A.
RC  STRAIN=22202v3.1;
RX  MEDLINE=99094949; PubMed=9878014;
RA  Van Dyke R.B., Korber B.T., Popek E., Macken C., Widmayer S.M., R.R.,
RA  Bardeguet A., Hansen I.C., Wiznia A., Luzuriaga K., Viscarello R.R.,
RA  Wolinsky S., the Ariel Core Investigators;
RT  "The Ariel Project: A prospective cohort study of maternal-child
RT  transmission of human immunodeficiency virus type 1 in the era of
RT  maternal antiretroviral therapy.";
RL  J. Infect. Dis. 179:319-328(1999).
DR  EMBL: AF112548; AAF13326.1; -.
FT  NON_TER 1
SQ  SEQUENCE 10 AA; 961 MW; 4EDD23640DD81DD7 CRC64;

Query Match 40.0%; Score 2; DB 15; Length 10;
Best Local Similarity 100.0%; Pred. No. 3 3e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
DB 1 GS 2

RESULT 46
P82181 ID P82181 PRELIMINARY; PRT; 6 AA.
AC P82181;
DT 01-JUN-2000 (TrEMBLrel. 14, Created)
DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Chloroplast 50S ribosomal protein L10 beta (Fragment).
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE.
RC STRAIN=cv. ALVARO; TISSUE=Leaf;
RX MEDLINE=20435798; PubMed=10874046;
RA Yamaguchi K., Subramanian A.R.;
RT "The plastid ribosomal proteins. Identification of all the proteins in
RT the 50 S subunit of an organelle ribosome (chloroplast).";
RL J. Biol. Chem. 275:28466-28482(2000).
CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC -!- MASS SPECTROMETRY: MW=16.5 KDA.
CC -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.
CC -!- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
DR InterPro; IPR001790; Ribosomal_L10.
DR Pfam; PF00466; Ribosomal_L10; PARTIAL.
DR PROSITE; PS01109; RIBOSOMAL_L10; PARTIAL.
KW Ribosomal protein; Chloroplast; rRNA-binding.
FT NON_TER 6
SQ SEQUENCE 6 AA; 675 MW; 6321B415B05DB000 CRC64;

Query Match 20.0%; Score 1; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 S 5
DB 3 S 3

RESULT 47
P82541 ID P82541 PRELIMINARY; PRT; 6 AA.
AC P82541;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

```

```

DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Chloroplast 30S ribosomal protein S19 beta (Fragment).
OS Spinacia oleracea (Spinach).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC STRAIN=cv. ALVARO; TISSUE=Leaf;
RX MEDLINE=20435797; PubMed=10874039;
RA Yamaguchi K., von Knoblauch K., Subramanian A.R.;
RT "The plastid ribosomal proteins. Identification of all the proteins in
RT the small subunit of an organelle ribosome (chloroplast).";
RL J. Biol. Chem. 275:28455-28465(2000).
CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 16S RIBOSOMAL RNA.
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC -!- MASS SPECTROMETRY: MW=10477.0; METHOD=ELECTROSPRAY.
CC -!- MASS SPECTROMETRY: MW=10495; METHOD=MALDI.
CC -!- MISCELLANEOUS: S19 ALPHA AND BETA FORMS DIFFER IN PI. S19 BETA
CC FORM IS THE MINOR BASIC FORM.
CC -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 12 KDA.
CC -!- SIMILARITY: BELONGS TO THE S19P FAMILY OF RIBOSOMAL PROTEINS.
DR InterPro; IPR002222; Ribosomal_S19.
DR Pfam; PF0203; Ribosomal_S19; PARTIAL.
DR PRINTS; PR00975; RIBOSOMAL_S19; PARTIAL.
DR PROSITE; PS00323; RIBOSOMAL_S19; PARTIAL.
KW Ribosomal protein; Chloroplast; rRNA-binding.
FT NON_TER 6
SQ SEQUENCE 6 AA; 732 MW; 63333735A411C000 CRC64;

Query Match 20.0%; Score 1; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 S 5
DB 3 S 3

RESULT 48
P82182 ID P82182 PRELIMINARY; PRT; 6 AA.
AC P82182;
DT 01-JUN-2000 (TrEMBLrel. 14, Created)
DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Chloroplast 50S ribosomal protein L10 gamma (Fragment).
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE.
RC STRAIN=cv. ALVARO; TISSUE=Leaf;
RX MEDLINE=20435798; PubMed=10874046;
RA Yamaguchi K., Subramanian A.R.;
RT "The plastid ribosomal proteins. Identification of all the proteins in
RT the 50 S subunit of an organelle ribosome (chloroplast).";
RL J. Biol. Chem. 275:28466-28482(2000).
CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.
CC -!- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
DR InterPro; IPR001790; Ribosomal_L10.
DR Pfam; PF00466; Ribosomal_L10; PARTIAL.
DR PROSITE; PS01109; RIBOSOMAL_L10; PARTIAL.
KW Ribosomal protein; Chloroplast; rRNA-binding.

```

```

FT SQ NON_TER 6 6 675 MW; 6321B415B05DB000 CRC64;
Query Match 20.0%; Score 1; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 S 5
DB 3 S 3

RESULT 49
Q8KMS3 PRELIMINARY; PRT; 7 AA.
ID Q8KMS3
AC Q8KMS3;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE Putative MerR2 protein.
GN MERR2
OS Klebsiella sp. LS13-39.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Klebsiella.
ON NCBI_TaxID=143776;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LS13-39;
RX MEDLINE=21604134; PubMed=11763342;
RA Mindlin S.Z., Kholodii G.Y., Gorlenko Z.M., Minakhina S.V.,
RA Minakhin L.S., Kalyaeva E.S., Kopteva A.V., Petrova M.A.,
RA Yuriyeva O.V., Nikiforov V.G.;
RT "Mercury resistance transposons of Gram-negative environmental
RT bacteria and their classification.";
RL Res. Microbiol. 152:811-822(2001).
DR EMBL; AJ302776; CAC82975.1; -.
SQ SEQUENCE 7 AA; 608 MW; 6DC1B5BDD87DD6F0 CRC64;

Query Match 20.0%; Score 1; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
DB 3 G 3

RESULT 50
Q47505 PRELIMINARY; PRT; 7 AA.
ID Q47505
AC Q47505;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE McclA protein.
GN McclA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
ON NCBI_TaxID=562;
RX [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=96099297; PubMed=8522520;
RA Gonzalez-Pastor J.B., San Millan J.L., Castilla M.A., Moreno F.;
RA "Structure and organization of plasmid genes required to produce the
RT translation inhibitor microcin C7.";
RL J. Bacteriol. 177:7131-7140(1995).
DR EMBL; X57583; CAA40808.1; -.
SQ SEQUENCE 7 AA; 763 MW; 644DD44861B406F0 CRC64;

Query Match 20.0%; Score 1; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

FT SQ NON_TER 6 6 675 MW; 6321B415B05DB000 CRC64;
Query Match 20.0%; Score 1; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 S 5
DB 3 S 3

RESULT 51
P70804 PRELIMINARY; PRT; 7 AA.
ID P70804
AC P70804;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DE Algt protein (Fragment).
GN ALGT.
OS Azotobacter vinelandii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Azotobacter.
ON NCBI_TaxID=354;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E;
RX MEDLINE=96427318; PubMed=8830682;
RA Rehm B.H.A., Ertesvag H., Valla S.;
RT "A new Azotobacter vinelandii mannuronan C-5-epimerase gene (algC) is
RT part of an alg gene cluster physically organized in a manner similar
RT to that in Pseudomonas aeruginosa.";
RL J. Bacteriol. 178:5884-5889(1996).
DR EMBL; X87973; CAA61230.1; -.
FT NON TER 1
SQ SEQUENCE 7 AA; 684 MW; 71B5A5A5A2D1AED0 CRC64;

Query Match 20.0%; Score 1; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 S 5
DB 4 S 4

RESULT 52
Q47029 PRELIMINARY; PRT; 7 AA.
ID Q47029
AC Q47029;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE Aad A1 protein (Fragment).
GN Aad A1.
OS Enterobacter cloacae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Enterobacter.
ON NCBI_TaxID=550;
RX [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=94079349; PubMed=8257126;
RA Rather P.N., Mann P.A., Mierzwa R., Hare R.S., Miller G.H., Shaw K.J.;
RT "Analysis of the aac(3)-Via gene encoding a novel 3-N-
RT acetyltransferase.";
RL Antimicrob. Agents Chemother. 37:2074-2079(1993).
DR EMBL; M88012; AAA16193.1; -.
FT NON TER 1
SQ SEQUENCE 7 AA; 744 MW; 633862D2C321A030 CRC64;

Query Match 20.0%; Score 1; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
DB 4 G 4

```

```

Db          6 G 6

RESULT 53
P72081
ID P72081      PRELIMINARY;      PRT;      7 AA.
AC P72081;
DT 01-FEB-1997 (TReMBLrel. 02, Created)
DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE 3'-methylcephem hydroxylase (Fragment).
GN CFF.
OS Nocardia lactamdurans.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Pseudonocardineae; Pseudonocardaceae; Amycolatopsis.
OX NCBI_TaxID=1913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96009872; PubMed=7557411;
RA Coque J., Perez-Liarena F.J., Enguita F.J., Fuente J.L., Martin J.F.,
RA Liras P.;
RT "Characterization of the cmh genes of Nocardia lactamdurans and
RT Streptomyces clavuligerus encoding a functional 3'-hydroxymethylcephem
RT O-carbamoyltransferase for cephamycin biosynthesis.";
RL Gene 162:21-27(1995).
DR EMBL; Z21682; CAA79797.1; -.
FT NON_TER 1
SQ SEQUENCE 7 AA; 746 MW; 75A1B2CDD1E681C0 CRC64;

Query Match      20.0%; Score 1; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 S 5
      |
Db      7 S 7

RESULT 54
O07354
ID O07354      PRELIMINARY;      PRT;      7 AA.
AC O07354;
DT 01-JUL-1997 (TReMBLrel. 04, Created)
DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE NifK (Fragment).
GN NIFK.
OS Synechococcus sp. (strain PCC 8801 / RF-1) (Cyanothecae PCC 8801).
OC Bacteria; Cyanobacteria; Chroococcales; Cyanothecae.
OX NCBI_TaxID=41431;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=RF-1;
RC MEDLINE=99231861; PubMed=10217509;
RA Huang T.C., Lin R.F., Chu M.K., Chen H.M.;
RT "Organization and expression of nitrogen-fixation genes in the aerobic
RT nitrogen-fixing unicellular cyanobacterium Synechococcus sp. strain
RT RF-1.";
RL Microbiology 145:743-753(1999).
DR EMBL; AF003700; AAC35193.1; -.
FT NON_TER 1
SQ SEQUENCE 7 AA; 849 MW; 7412C72AA9D5B030 CRC64;

Query Match      20.0%; Score 1; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 S 5
      |
Db      2 S 2

RESULT 55

```

```

Q8MFY6
ID Q8MFY6      PRELIMINARY;      PRT;      7 AA.
AC Q8MFY6;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE PsbA (Fragment).
GN PsbA.
OS Taraxacum (sect. Dioszegia) sp. 4310Hnew.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; campanulids; Asterales; Asteraceae; Cichorioideae;
OC Cichorieae; Taraxacum.
OX NCBI_TaxID=154248;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RA Mes T.H.M.;
RT "Reconstruction of the evolution of trnf pseudogenes.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY015477; AAK21591.1; -.
KW Chloroplast.
FT NON_TER 1
SQ SEQUENCE 7 AA; 675 MW; 687451BSA76DDB70 CRC64;

Query Match      20.0%; Score 1; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 S 5
      |
Db      4 S 4

RESULT 56
Q95945
ID Q95945      PRELIMINARY;      PRT;      7 AA.
AC Q95945;
DT 01-FEB-1997 (TReMBLrel. 02, Created)
DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Inside intron 5 (Fragment).
OS Saccharomyces cerevisiae (Baker's yeast).
OG Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=D273-10B;
RC MEDLINE=81069885; PubMed=6254986;
RA Bonitz S.G., Coruzzi G., Thalenfeld B., Tzagoloff A., Macino G.;
RT "Assembly of the mitochondrial membrane system: Structure and
RT nucleotide sequence of the gene coding for subunit 1 of yeast
RT cytochrome oxidase.";
RL J. Biol. Chem. 255:11927-11941(1980).
DR EMBL; V00694; CAA24066.1; -.
KW Mitochondrion.
FT NON_TER 1
SQ SEQUENCE 7 AA; 859 MW; 75B7232362CDC460 CRC64;

Query Match      20.0%; Score 1; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 S 5
      |
Db      7 S 7

RESULT 57
P93233

```

ID P93233 PRELIMINARY; PRT; 7 AA.
 AC P93233;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE 1-aminocyclopropane-1-carboxylate synthase (EC 4.4.1.14)
 DE (Fragment).
 DE LE-ACS1B.
 GN GN
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; lamiids; Solanales; Solanaceae; Solanum.
 ON NCBI_TaxID=4081;
 RX [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=97351561; PubMed=9207843;
 RA Oetiker J.H., Olson D.C., Shiu O.Y., Yang S.F.;
 RT "Differential induction of seven 1-aminocyclopropane-1-carboxylate
 RT synthase genes by elicitor in suspension cultures of tomato
 RT (Lycopersicon esculentum).";
 RL Plant Mol. Biol. 34:275-286(1997).
 DR EMBL; U75692; AAC49682.1; -.
 KW Lyase.
 FT NON_TER 1 1
 SQ SEQUENCE 7 AA; 828 MW; 71B412C7377415D0 CRC64;

Query Match 20.0%; Score 1; DB 10; Length 7;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 S 5
 |
 Db 1 S 1

RESULT 58
 O49223 PRELIMINARY; PRT; 7 AA.
 ID O49223;
 AC O49223;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE HMG-1-like protein (Fragment).
 DE Glycine max (Soybean).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 ON NCBI_TaxID=3847;
 RX [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=cv. Essex; TISSUE=Root;
 RC MEDLINE=91367679; PubMed=1891369;
 RA Laux T., Goldberg R.B.;
 RT "A plant DNA binding protein shares highly conserved sequence motifs
 RT with HMG-box proteins.";
 RL Nucleic Acids Res. 19:4769-4769(1991).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=cv. Essex; TISSUE=Root;
 RC MEDLINE=91367679; PubMed=1891369;
 RA Mahalingam R., Knap H.T.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF047050; AAC03556.1; -.
 FT NON_TER 1 1
 SQ SEQUENCE 7 AA; 850 MW; 6AAAAAB378637810 CRC64;

Query Match 20.0%; Score 1; DB 10; Length 7;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 |
 Db 1 G 1

RESULT 59
 Q9C5B3 PRELIMINARY; PRT; 7 AA.
 ID Q9C5B3;
 AC Q9C5B3;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hypothetical 0.7 kDa protein (Fragment).
 GN DIB1 10A-2B.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 ON NCBI_TaxID=3702;
 RX [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Root;
 RC MEDLINE=21171025; PubMed=11277426;
 RA Vercauteren I., Van Der Schueren E., Van Montagu M., Gheysen G.;
 RT "Arabidopsis thaliana genes expressed in the early compatible
 RT interaction with root-knot nematodes.";
 RL Mol. Plant Microbe Interact. 14:288-299(2001).
 DR EMBL; AJ286350; CAB71014.2; -.
 KW Hypothetical protein.
 FT NON_TER 1 1
 FT NON_TER 7 7
 SQ SEQUENCE 7 AA; 719 MW; 6732C7287EB325D0 CRC64;

Query Match 20.0%; Score 1; DB 10; Length 7;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 S 5
 |
 Db 1 S 1

RESULT 60
 P82445 PRELIMINARY; PRT; 7 AA.
 ID P82445;
 AC P82445;
 DT 01-JUN-2000 (TrEMBLrel. 14, Created)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE 10 kDa cell wall protein (fragment).
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
 ON NCBI_TaxID=4097;
 RX [1]
 RN SEQUENCE.
 RC STRAIN=cv. PETIT HAVANA;
 RA Blee K.A., Bonham V.A., Mitchell G.P., Robertson D., Slabas A.R.,
 RA Wojtaszek P., Bolwell G.P.;
 RT "Proteomic study of secondary cell wall proteins from transformed
 RT tobacco culture.";
 RL Planta 0:0-0(2000).
 CC -!- SUBCELLULAR LOCATION: CELL WALL.
 CC -!- TISSUE SPECIFICITY: XYLEM.
 KW Cell wall.
 FT NON_TER 7 7
 SQ SEQUENCE 7 AA; 758 MW; 69D2C1E862D1B2A0 CRC64;

Query Match 20.0%; Score 1; DB 10; Length 7;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 |
 Db 4 G 4

```

RESULT 61
Q63668      PRELIMINARY;      PRT;      7 AA.
ID Q63668
AC Q63668;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ORF2 protein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC SPRAIN=Sprague-Dawley; TISSUE=Kidney;
RX MEDLINE=9539650; PubMed=7667072;
RA Mandon B., Bellanger A.C., Elalouf J.M.;
RT "Inverse-PCR-mediated cloning of the promoter for the rat vasopressin
RT V2 receptor gene.";
RL Pflugers Arch. 430:12-18(1995).
DR EMBL; X83264; CAA58237.1; -
SQ SEQUENCE 7 AA; 703 MW; 75A767287DC6D6F0 CRC64;

Query Match      20.0%; Score 1; DB 11; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
DB 4 G 4

RESULT 62
Q65578      PRELIMINARY;      PRT;      7 AA.
ID Q65578
AC Q65578;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 0.8 kDa protein (Fragment).
OS Bovine herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10320;
RN [1]
RP SEQUENCE FROM N.A.
RC SPRAIN=Cooper;
RX MEDLINE=95313343; PubMed=7793062;
RA Vlcek C., Benes V., Lu Z., Kutish G.F., Paces V., Rock D.,
RA Letchworth G.J., Schwytzer M.;
RT "Nucleotide sequence analysis of a 30-kb region of the bovine
RT herpesvirus 1 genome which exhibits a colinear gene arrangement with
RT the UL21 to UL4 genes of herpes simplex virus.";
RL Virology 210:100-108(1995).
DR EMBL; Z48053; CAA88130.1; -
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 7 AA; 758 MW; 6DD33455B1F1B1C0 CRC64;

Query Match      20.0%; Score 1; DB 12; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 S 5
DB 4 S 4

RESULT 63
Q8JJ20      PRELIMINARY;      PRT;      7 AA.
ID Q8JJ20
AC Q8JJ20;

```

```

DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Extracellular fatty acid binding protein (Fragment).
GN EXPFABP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RT Wang O., Li N., Li H.;
RT "Cloning and sequencing of 3' UTR of EXPFABP gene in chicken.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF487519; AAL96665.1; -
FT NON TER 1
SQ SEQUENCE 7 AA; 780 MW; 72CB1AB2D5BEBB70 CRC64;

Query Match      20.0%; Score 1; DB 13; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 S 5
DB 3 S 3

RESULT 64
Q42564      PRELIMINARY;      PRT;      7 AA.
ID Q42564
AC Q42564;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE Truncated voltage-gated sodium channel alpha subunit (Fragment).
GN SCN8A.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97442476; PubMed=9295353;
RA Plummer N.W., McBurney M.W., Meisler M.H.;
RT "Alternative splicing of the sodium channel SCN8A predicts a truncated
RT two-domain protein in fetal brain and non-neuronal cells.";
RL J. Biol. Chem. 272:24008-24015(1997).
DR EMBL; U97673; AAB80916.1; -
KW Ionic channel.
FT NON TER 1
SQ SEQUENCE 7 AA; 730 MW; 75B72EA2C73772A0 CRC64;

Query Match      20.0%; Score 1; DB 13; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 S 5
DB 7 S 7

RESULT 65
Q07624      PRELIMINARY;      PRT;      7 AA.
ID Q07624
AC Q07624;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE UORFL.

```

OS Rous sarcoma virus (strain Prague C).
OC Viruses; Retroid viruses; Retroviridae; Alpharetrovirus.
OXX NCBI_TaxID=11888;
RN [1]
SEQUENCE FROM N.A.
Q9AGP4 MEDLINE=93010967; PubMed=1327749;
AC Donze O., Spahr P.F.;
ID "Role of the open reading frames of Rous sarcoma virus leader RNA in
RT translation and genome packaging.";
RT EMBO J. 11:3747-3757(1992).
RL EMBL; X67587; CAA47862.1; -;
RS Q9AGP4 SEQUENCE 7 AA; 672 MW; 776045A7687DD6F0 CRC64;
SQ
Query Match 20.0%; Score 1; DB 15; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 G 1
DB 3 G 3
RESULT 66
Q9AGP4 PRELIMINARY; PRT; 8 AA.
AC Q9AGP4:
ID 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-WAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Serine hydroxymethyltransferase (Fragment).
OS GLYA.
OC Arthrobacter sp. 11N.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococcineae; Micrococcaceae; Arthrobacter.
OXX NCBI_TaxID=153502;
RN [1]
SEQUENCE FROM N.A.
RT STRAIN=11N;
RA Meskys R., Harris R.J., Caeate V., Baaran J., Scrutton N.S.;
RT "Genetic organization of the genes involved in dimethylglycine and
RT sarcosine degradation in Arthrobacter spp.: implications for glycine
RT betaine catabolism";
RT Submitted (DDBJ-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329478; AAK16486.1; -;
KW Methyltransferase; Transferase.
FT NON_TER 1
SQ SEQUENCE 8 AA; 898 MW; 6B18705333372457 CRC64;
Query Match 20.0%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 G 1
DB 7 G 7
RESULT 67
Q9R3X0 PRELIMINARY; PRT; 8 AA.
ID Q9R3X0:
AC Q9R3X0:
ID 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 16, Last annotation update)
DE Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit
DE (fragment).
OS RBCL.
GN RBCL.
OC Planktothrix rubescens.
OC Bacteria; Cyanobacteria; Oscillatoriales; Planktothrix.
OXX NCBI_TaxID=59512;
RN [1]
SEQUENCE FROM N.A.
RT STRAIN=BC-PLA 9316, and BC-PLA 9303;
RC

RX MEDLINE=20005589; PubMed=10537197;
RA Beard S.J., Handley B.A., Hayes P.K., Walsby A.E.;
RT "The diversity of gas vesicle genes in Planktothrix rubescens from
RT Lake Zurich.";
RL Microbiology 145:2757-2768(1999).
DR EMBL; AJ132249; CAB59537.1; -;
DR EMBL; AJ132248; CAB59534.1; -;
FT NON_TER 1
SQ SEQUENCE 8 AA; 957 MW; 33D1AAA685BB19CB CRC64;
Query Match 20.0%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 5 S 5
DB 4 S 4
RESULT 68
Q9RSR2 PRELIMINARY; PRT; 8 AA.
ID Q9RSR2:
AC Q9RSR2:
ID 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 31,000 DA product of ORF8 (Fragment).
OS Shigella dysenteriae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OXX NCBI_TaxID=622;
RN [1]
SEQUENCE
RX MEDLINE=92085268; PubMed=1660923;
RA Polard P., Preere M.F., Chandler M., Fayet O.;
RT "Programmed translational frameshifting and initiation at an AUU codon
RT in gene expression of bacterial insertion sequence IS911.";
RL J. Mol. Biol. 222:465-477(1991).
FT NON_TER 8
SQ SEQUENCE 8 AA; 902 MW; FE2DCAF8586AE336 CRC64;
Query Match 20.0%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 G 1
DB 4 G 4
RESULT 69
O68485 PRELIMINARY; PRT; 8 AA.
ID O68485:
AC O68485:
ID 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Aminoglycoside 3'-O-nucleotidyltransferase (Fragment).
GN AADAI.
OS Klebsiella pneumoniae.
OG Plasmid pQ1000.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Klebsiella.
OXX NCBI_TaxID=573;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=KI;
RX MEDLINE=98287600; PubMed=9624504;
RA Centron D., Roy P.H.;
RT "Characterization of the 6'-N-aminoglycoside acetyltransferase gene
RT aac(6')-Iq from the integron of a natural multiresistance plasmid.";
RL Antimicrob. Agents Chemother. 42:1506-1508(1998).
DR EMBL; AF047556; AAC25501.1; -;

KW Plasmid; Transferase.
FT NON TER 8 8
SQ SEQUENCE 8 AA; 878 MW; EFC1B5A2CB1DD056 CRC64;

Query Match 20.0%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 S 5
|
Db 6 S 6

RESULT 70
Q93SR0 PRELIMINARY; PRT; 8 AA.
AC Q93SR0; 201 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Beta-lactamase repressor Blai (Fragment).
GN BLAI.
OS Staphylococcus epidermidis.
OG Plasmid pST6.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=6;
RA Sidhu M.S., Heir E., Sorum H., Holck A.L.;
RT "Genetic linkage between quaternary ammonium compound and beta-lactam resistance in Staphylococci isolated from food."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY028779; AAK38453.1; -.
KW Plasmid.
FT NON TER 1
SQ SEQUENCE 8 AA; 930 MW; 4E3325B05AA44720 CRC64;

Query Match 20.0%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 S 5
|
Db 6 S 6

RESULT 71
Q8RJ0 PRELIMINARY; PRT; 8 AA.
AC Q8RJ0; 21 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE DNA replication initiation protein (Fragment).
GN DNAA-Like.
OS Streptomyces caespitosus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC27422;
RA Ma W., Mao X., Jie L., Qin Z.J., Jiang W.H., Chiao R.S., Zhao G.P.;
RT "Cloning, sequence analysis and function analysis of the replication origin oric of Streptovericillum caespitosus ATCC27422."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ458440; CAD30324.1; -.
FT NON TER 1
SQ SEQUENCE 8 AA; 915 MW; 04744330440451A7 CRC64;

Query Match 20.0%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
|
Db 8 G 8

RESULT 72
Q44463 PRELIMINARY; PRT; 8 AA.
AC Q44463; 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Agrobacterium tumefaciens Ti plasmid virD3 and virD4 genes (Fragment).
OS Rhizobiales (rhizobacteria).
OG Plasmid Ti.
OC Bacteria; Proteobacteria; Alphaproteobacteria.
OX NCBI_TaxID=356;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88015611; PubMed=3658701;
RA Portner S.G., Yanofsky M.F., Nester B.W.;
RT "molecular characterization of the virD operon from Agrobacterium tumefaciens."
RL Nucleic Acids Res. 15:7503-7517(1987).
DR EMBL: X06045; CAA29439.1; -.
KW Plasmid.
FT NON TER 8
SQ SEQUENCE 8 AA; 887 MW; F8F2C325B33861A6 CRC64;

Query Match 20.0%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
|
Db 3 G 3

RESULT 73
P77556 PRELIMINARY; PRT; 8 AA.
AC P77556; 02 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DE TRAY (Fragment).
GN TRAY.
OS Escherichia coli.
OG Plasmid IncFII R1.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ECOR11;
RX MEDLINE=96400508; PubMed=8807284;
RA Boyd E.F., Hill C.W., Rich S.W., Hartl D.L.;
RT "Mosaic structure of plasmids from natural populations of Escherichia coli."
RL Genetics 143:1091-1100(1996).
DR EMBL: U50661; AAC44245.1; -.
DR EMBL: U50850; AAC44234.1; -.
DR EMBL: U50851; AAC44235.1; -.
DR EMBL: U50852; AAC44236.1; -.
DR EMBL: U50853; AAC44237.1; -.
DR EMBL: U50654; AAC44238.1; -.
DR EMBL: U50655; AAC44239.1; -.
DR EMBL: U50856; AAC44240.1; -.
DR EMBL: U50657; AAC44241.1; -.
DR EMBL: U50658; AAC44242.1; -.

DR EMBL; U50659; AAC44243.1; -;
 DR EMBL; U50660; AAC44244.1; -;
 KW Plasmid.
 FT NON_TER 1 1
 SQ SEQUENCE 8 AA; 834 MW; D335A5B0544735A1 CRC64;
 Query Match 20.0%; Score 1; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 S 5
 Db 2 S 2
 RESULT 74
 Q9ZE29 PRELIMINARY; PRT; 8 AA.
 ID Q9ZE29
 AC Q9ZE29
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
 DE 2-isopropylmalate synthase (EC 4.1.3.12) (Fragment).
 LEUA.
 GN Buchnera aphidicola.
 OS Buchnera aphidicola.
 OC Plasmid pBrc1.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Buchnera.
 OX NCBI_TaxID=9;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=99028904; PubMed=9812361;
 RX Silva F.J., Van Ham R.C.H.J., Sabater B., Latorre A.;
 RT "Structure and evolution of the leucine plasmids carried by the
 RT endosymbiont (Buchnera aphidicola) from aphids of the family
 RT Aphididae."
 RL FEMS Microbiol. Lett. 168:43-49 (1998).
 DR EMBL; AJ006874; CAA07290.1; -;
 KW Lyase; Plasmid.
 FT NON_TER 8 8
 SQ SEQUENCE 8 AA; 917 MW; FE205042D6C5B446 CRC64;
 Query Match 20.0%; Score 1; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 S 5
 Db 3 S 3
 RESULT 75
 O09258 PRELIMINARY; PRT; 8 AA.
 ID O09258
 AC O09258
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE NifH (Fragment).
 DE NifH.
 GN Synechococcus sp. (strain PCC 8801 / RP-1) (Cyanothecae PCC 8801).
 OS Synechococcus; Cyanobacteria; Chroococcales; Cyanothecae.
 OC NCBI_TaxID=41431;
 OX [1]
 RP SEQUENCE FROM N.A.
 RN STRAIN=RP-1;
 RX MEDLINE=99231861; PubMed=10217509;
 RA Huang T.C., Lin R.P., Chu M.K., Chen H.M.;
 RT "Organization and expression of nitrogen-fixation genes in the aerobic
 RT nitrogen-fixing unicellular cyanobacterium Synechococcus sp. strain
 RT RP-1."
 RL Microbiology 145:743-753 (1999).
 DR EMBL; AF001780; AAC33369.1; -;

FT NON_TER 8 8
 SQ SEQUENCE 8 AA; 985 MW; F16B59CDD046C406 CRC64;
 Query Match 20.0%; Score 1; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 G 1
 Db 8 G 8
 RESULT 76
 O8RSR3 PRELIMINARY; PRT; 8 AA.
 ID O8RSR3
 AC O8RSR3
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE PduG protein (Fragment).
 GN PDUG.
 OS Lactobacillus collinoides.
 OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
 OC Lactobacillus.
 OX NCBI_TaxID=33960;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LMG 18850;
 RA Sauvageot N., Gouffi K., Laplace J.M., Auffray Y.;
 RT "Characterisation of the diol dehydratase pdu operon of Lactobacillus
 RT collinoides."
 RL FEMS Microbiol. Lett. 0:0-0(0).
 DR EMBL; AJ297723; CAD01093.1; -;
 FT NON_TER 8 8
 SQ SEQUENCE 8 AA; 982 MW; ECA40B05B8B5DD6 CRC64;
 Query Match 20.0%; Score 1; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 S 5
 Db 3 S 3
 RESULT 77
 Q9X3K1 PRELIMINARY; PRT; 8 AA.
 ID Q9X3K1
 AC Q9X3K1
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
 DE Cytochrome b (Fragment).
 DE PETB.
 OS Prochlorococcus sp.
 OS Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
 OC Prochlorococcus.
 OX NCBI_TaxID=1220;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Urbach E., Chisholm S.W.;
 RT "Genetic diversity in Prochlorococcus populations flow cytometrically
 RT sorted from the Sargasso Sea and Gulf Stream."
 RL Limnol. Oceanogr. 43:1615-1630 (1998).
 DR EMBL; AF070193; AAD23233.1; -;
 FT NON_TER 1 1
 SQ SEQUENCE 8 AA; 799 MW; 10376865B72866D3 CRC64;
 Query Match 20.0%; Score 1; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 G 1

```

Db          3 G 3

RESULT 78
O87471      PRELIMINARY;      PRT;      8 AA.
ID O87471;
AC O87471;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE Hifa (Fragment).
GN HIFA.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Eagan;
RX MEDLINE=98389689; PubMed=97211313;
RA Mlanga-Mutangadura T., Morlin G., Smith A.L., Eisenstark A.,
RA Golomb M.;
RT "Evolution of the major pilus gene cluster of haemophilus
RT influenzae.";
RL J. Bacteriol. 180:4693-4703(1998).
DR EMBL; AF071762; AAC35830.1; -.
FT NON TER 1
SQ SEQUENCE 8 AA; 876 MW; DAB44451A7272325 CRC64;

Query Match      20.0%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          5 S 5
Db          1 S 1

RESULT 79
Q56140      PRELIMINARY;      PRT;      8 AA.
ID Q56140;
AC Q56140;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE STP6 protein (Fragment).
GN STP6.
OS Streptococcus thermophilus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1308;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Still;
RX MEDLINE=95047254; PubMed=7958782;
RA Constable A., Mollet B.;
RT "Isolation and characterisation of promoter regions from Streptococcus
RT thermophilus.";
RL FEMS Microbiol. Lett. 122:85-90(1994).
DR EMBL; X78210; CAA55045.1; -.
FT NON TER 8
SQ SEQUENCE 8 AA; 846 MW; ED086772D5B045B6 CRC64;

Query Match      20.0%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          1 G 1
Db          7 G 7

RESULT 80
Q56246      PRELIMINARY;      PRT;      8 AA.
ID Q56246;
AC Q56246; Q52410;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE Thermophilic proton ATPase epsilon subunit amino terminal (Fragment).
OS Thermophilic bacterium PS3.
OC Bacteria; Firmicutes; Bacilliales.
OX NCBI_TaxID=2334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PS3;
RX MEDLINE=87137359; PubMed=2880841;
RA Kagawa Y., Ishizuka M., Saishu T., Nakao S.;
RT "Stable structure of thermophilic proton ATPase beta subunit.";
RL J. Biochem. 100:923-934(1986).
DR EMBL; D00113; BAA00067.1; -.
DR EMBL; X04609; CAA28278.1; -.
FT NON TER 8
SQ SEQUENCE 8 AA; 914 MW; EBD5A2C1E041A336 CRC64;

Query Match      20.0%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          5 S 5
Db          7 S 7

RESULT 81
Q56429      PRELIMINARY;      PRT;      8 AA.
ID Q56429;
AC Q56429;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CAPDH (Fragment).
OS Thermus thermophilus.
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=274;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HB-8;
RX MEDLINE=89025722; PubMed=3052437;
RA Bowen D., Littlechild J.A., Fothergill J.E., Watson H.C., Hall L.;
RT "Nucleotide sequence of the phosphoglycerate kinase gene from the
RT extreme thermophile, Thermus thermophilus.";
RL Biochem. J. 254:509-517(1988).
DR EMBL; X12464; CAA31005.1; -.
FT NON TER 1
SQ SEQUENCE 8 AA; 885 MW; 33C873333732C72B CRC64;

Query Match      20.0%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          1 G 1
Db          7 G 7

RESULT 82
Q9R9E0      PRELIMINARY;      PRT;      8 AA.
ID Q9R9E0;
AC Q9R9E0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE Stage V sporulation protein E (Fragment).

```

GN SPOVE.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=93003529; PubMed=1391053;
RA Henriques A.O., de Lencastre H., Pigott P.J.;
RT "A Bacillus subtilis morphogene cluster that includes spore is
homologous to the mra region of Escherichia coli.";
RL Biochimie 74:735-748(1992).
RR EMBL; X64258; CAA45556.1; -.
FT NON_TER 8
SQ SEQUENCE 8 AA; 893 MW; EE75A1A33321B1A6 CRC64;

Query Match 20.0%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 S 5
Db 7 S 7

RESULT 83
Q93SP2 PRELIMINARY; PRT; 8 AA.
ID Q93SP2
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Quaternary ammonium compound-resistance protein QacEdelta1
(Fragment).
DE QACEDelta1.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TYMC704; TRANSPOSON=class I integron;
RA Lee K., Chong Y., Yum J.H., Yong D., Livermore D.M.;
RT "VIM-2 metallo-beta-lactamase gene-containing integron in a
Pseudomonas aeruginosa clinical isolate.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RR EMBL; AY029772; AAK50441.1; -.
FT NON_TER 8
SQ SEQUENCE 8 AA; 920 MW; FE2729C732C87336 CRC64;

Query Match 20.0%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 G 1
Db 3 G 3

RESULT 84
Q45889 PRELIMINARY; PRT; 8 AA.
ID Q45889
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HA-11 protein (Fragment).
DE HA-11.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=667AB;
RX MEDLINE=96210012; PubMed=8631890;
RA Hutson R.A., Zhou Y., Collins M.D., Johnson E.A., Hatheway C.L.,
RA Sugiyama H.;
RT "Genetic Characterisation of Clostridium botulinum type A containing
silent type B neurotoxin gene sequences.";
RL J. Biol. Chem. 271:10786-10792(1996).
RR EMBL; X87850; CAA61130.1; -.
FT NON_TER 8
SQ SEQUENCE 8 AA; 982 MW; FE29D1B40B02D5B6 CRC64;

Query Match 20.0%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 S 5
Db 2 S 2

RESULT 85
Q9RSR0 PRELIMINARY; PRT; 8 AA.
ID Q9RSR0
AC Q9RSR0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 11,500 DA product of ORFA (Fragment).
OS Shigella dysenteriae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=622;
RN [1]
RP SEQUENCE.
RX MEDLINE=92085268; PubMed=1660923;
RA Polard P., Prere M.F., Chandler M., Fayet O.;
RT "Programmed translational frameshifting and initiation at an AUU codon
in gene expression of bacterial insertion sequence IS911.";
RL J. Mol. Biol. 222:465-477(1991).
FT NON_TER 8
SQ SEQUENCE 8 AA; 935 MW; F4C5B9C45AE33336 CRC64;

Query Match 20.0%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 S 5
Db 7 S 7

RESULT 86
Q9RC2 PRELIMINARY; PRT; 8 AA.
ID Q9RC2
AC Q9RC2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Plasmid cp32-3, possible partition proteins (Fragment).
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid cp32-3.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B31;

RX MEDLINE=98361033; PubMed=9695920;
RA Stevenson B., Casjens S., Rosa P.;
RT "Evidence of past recombination events among the genes encoding the
Erp antigens of Borrelia burgdorferi.";
RL Microbiology 144:1869-1879(1998).
RR EMBL; AF022480; AAC35445.1; -.

```

KW Plasmid.
FT NON TER
SQ SEQUENCE 8 AA; 985 MW; E8841B41A735B446 CRC64;

Query Match 20.0%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 S 5
Db 3 S 3

RESULT 87
Q47273 ID Q47273 PRELIMINARY; PRT; 8 AA.
AC Q47273;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Orf33, orf151, orf56, orf96, rus, orf45, orf8, orf127, and nmpC genes.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K-12;
RA MEDLINE=96196428; PubMed=8648624;
RA Mahdi A.A., Sharples G.J., Mandal T.N., Lloyd R.G.;
RT "Holliday Junction Resolvases encoded by homologous rusa genes in
RT Escherichia coli K-12 and Phage 32.";
RL J. Mol. Biol. 257:561-573(1996).
DR EMBL; X92587; CAA63323.1; -
SQ SEQUENCE 8 AA; 935 MW; FD4041A05B05B336 CRC64;

Query Match 20.0%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 S 5
Db 3 S 3

RESULT 88
Q9R4M3 ID Q9R4M3 PRELIMINARY; PRT; 8 AA.
AC Q9R4M3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE C0B1=BACTERIAL sex pheromone.
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE.
RC MEDLINE=95290767; PubMed=7772836;
RA Nakayama J., Abe Y., Ono Y., Isegai A., Suzuki A.;
RT "Isolation and structure of the Enterococcus faecalis sex pheromone,
RT C0B1, that induces conjugal transfer of the hemolysin/bacteriocin
RT plasmids, p0B1 and pY11.";
RL Biosci. Biotechnol. Biochem. 59:703-705(1995).
SQ SEQUENCE 8 AA; 741 MW; 83D87732C732CDC2 CRC64;

Query Match 20.0%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 G 1
Db 1 G 1

RESULT 89
Q934S4 ID Q934S4 PRELIMINARY; PRT; 8 AA.
AC Q934S4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MerD protein (Fragment).
GN MERD.
OS Thiobacillus ferrooxidans.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Acidithiobacillus.
OX NCBI_TaxID=920;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=G66; TRANSPOSON=Tn5037;
RA Kalyaeva E.S., Kholodii G.Y., Bass I.A., Gorlenko A.M., Yurieva O.V.,
RA Nikiforov V.G.;
RT "Tn5037, a Tn21-like mercury resistance transposon from Thiobacillus
RT ferrooxidans.";
RL Russ. J. Genet. 37:972-975(2001).
DR EMBL; AJ251743; CAC69252.1; -
FT NON TER
SQ SEQUENCE 8 AA; 937 MW; ED15A2D77B5DD446 CRC64;

Query Match 20.0%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 S 5
Db 7 S 7

RESULT 90
P72221 ID P72221 PRELIMINARY; PRT; 8 AA.
AC P72221;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Alginatase lyase (Fragment).
GN ALY.
OS Pseudomonas sp. (strain OS-ALG-9).
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=86038;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OS-ALG-9;
RA Fujiyama K.;
RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=OS-ALG-9;
RX MEDLINE=93329366; PubMed=8336113;
RA Maki H., Mori A., Fujiyama K., Kinoshita S., Yoshida T.;
RT "Cloning, sequence analysis and expression in Escherichia coli of a
RT gene encoding an alginatase lyase from Pseudomonas sp. OS-ALG-9.";
RL J. Gen. Microbiol. 139:987-993(1993).
DR EMBL; D38469; BAA21704.1; -
GN Lyase.
FT NON TER
SQ SEQUENCE 8 AA; 841 MW; 461DDDC5A5B041BB CRC64;

Query Match 20.0%; Score 1; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 S 5
Db 5 S 5

```

RA Coolbaugh M.I., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,

RA Caskey C.T.H.;

RT "Isolation of chromosome-specific genes by reciprocal probing of

RT arrayed cDNAs and cosmid libraries.";

RL Hum. Mol. Genet. 0:0-0(1995).

DR EMBL; L32074; AAA73884.1; --

FT NON_TER 1 1

FT NON_TER 8 8

SQ SEQUENCE 8 AA; 952 MW; EEC735B1E1F1B6D6 CRC64;

Query Match 20.0%; Score 1; DB 4; Length 8;

Best Local Similarity 100.0%; Pred. No. 8.3e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 S 5

Db 6 S 6

RESULT 100

Q16428

ID Q16428 PRELIMINARY; PRT; 8 AA.

AC Q16428;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)

DE Dystrophin protein (Fragment).

GN DYSTROPHIN.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=96163501; PubMed=8566960;

RA Holder E., Maeda M., Bies R.D.;

RT "Expression and regulation of the dystrophin Purkinje promoter in

RT human skeletal muscle, heart, and brain.";

RL Hum. Genet. 97:232-239(1996).

RN [2]

RP SEQUENCE FROM N.A.

RA Kitano T., Kobayakawa H., Saitou N.;

RT "Silver Project.";

RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; S81419; AAD14363.1; --

DR EMBL; AB037493; BAA90413.1; --

FT NON_TER 8 8

SQ SEQUENCE 8 AA; 825 MW; F21865A5A2CB05B6 CRC64;

Query Match 20.0%; Score 1; DB 4; Length 8;

Best Local Similarity 100.0%; Pred. No. 8.3e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 G 1

Db 7 G 7

Search completed: October 21, 2003, 18:27:13

Job time : 42.7222 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:12:30 ; Search time 8.61111 Seconds
(without alignments)
27.306 Million cell updates/sec

Title: US-10-057-890A-16

Perfect score: 5

Sequence: 1 GGGGS 5

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 372

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Listing first 150 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3	60.0	9	1	THYF_PIG
2	2	40.0	4	1	EOSI_HUMAN
3	2	40.0	4	1	OCPI_OCTMI
4	2	40.0	6	1	CIP1_MYTED
5	2	40.0	7	1	FAR5_HIRME
6	2	40.0	7	1	LANC_CARUI
7	2	40.0	8	1	ALL6_CARMA
8	2	40.0	8	1	ALL9_CARMA
9	2	40.0	8	1	LCK5_LEUMA
10	2	40.0	8	1	UF06_MOUSE
11	2	40.0	8	1	VGLG_HSV2B
12	2	40.0	8	1	WPI_PERAT
13	2	40.0	9	1	D1_NEPNO
14	2	40.0	9	1	DSIP_RABIT
15	2	40.0	9	1	FAR6_MACRS
16	2	40.0	9	1	FIBB_EYRPA
17	2	40.0	9	1	FIBB_THEGE
18	2	40.0	9	1	NEF_HY128
19	2	40.0	9	1	XYLA_STRSQ
20	2	40.0	10	1	AMPN_HELAM
21	2	40.0	10	1	CXHX_ONCMY
22	2	40.0	10	1	CU30_LOCM1
23	2	40.0	10	1	ESTA_SCHGA
24	2	40.0	10	1	GRP_RANR1
25	2	40.0	10	1	NO40_TOBAC
26	2	40.0	10	1	TK02_BOVIN
27	2	40.0	10	1	TKU1_UREUN
28	2	40.0	10	1	TPIS_NICPL
29	2	40.0	10	1	UPAB_HUMAN
30	1	20.0	3	1	GRWM_HUMAN
31	1	20.0	4	1	ACH1_ACHFU
32	1	20.0	4	1	DCML_PSECH
33	1	20.0	4	1	OCPI_OCTMI

34	1	20.0	4	1	RM01_YEAST	P36515	saccharomyc
35	1	20.0	5	1	AL14_CARMA	P81817	carcinus ma
36	1	20.0	5	1	BIOB_CITFR	P12997	citrobacter
37	1	20.0	5	1	PA22_PARMA	P81864	pardachirus
38	1	20.0	5	1	RE32_LITRU	P82073	litorea rub
39	1	20.0	5	1	TPIS_CANFA	P54714	canis famli
40	1	20.0	5	1	UF01_MOUSE	P38639	mus musculus
41	1	20.0	5	1	UXA4_CHLTR	P38005	chlamydia t
42	1	20.0	6	1	CIP2_MYTED	P13737	mytilus edu
43	1	20.0	6	1	FARP_MONEX	P41966	moniezia ex
44	1	20.0	6	1	LOK1_LOCM1	P41491	locusta mig
45	1	20.0	7	1	ALL2_CARMA	P81805	carcinus ma
46	1	20.0	7	1	ALL3_CARMA	P81806	carcinus ma
47	1	20.0	7	1	ALL4_CARMA	P81807	carcinus ma
48	1	20.0	7	1	ALL5_CARMA	P81808	carcinus ma
49	1	20.0	7	1	ALL7_CVDPO	P82158	cydia pomon
50	1	20.0	7	1	BRHP_CONIM	P58803	conus imper
51	1	20.0	7	1	CHOX_ALCSP	P61601	alcaligenes
52	1	20.0	7	1	CIA_ENTFA	P11932	enterococcu
53	1	20.0	7	1	FAR1_HELTI	P41871	helisoma tr
54	1	20.0	7	1	FAR3_HAECO	P81298	haemomachus
55	1	20.0	7	1	FAR3_PANRE	P41874	panagrellus
56	1	20.0	7	1	GPRP_MOUSE	P99025	mus musculus
57	1	20.0	7	1	HY7_PIG	P01153	sus scrofa
58	1	20.0	7	1	IGAO_DACDE	P06294	dactylium d
59	1	20.0	7	1	MNP1_LEPDE	P42984	leptinotars
60	1	20.0	7	1	PPH2_LVCES	P83379	lycopersico
61	1	20.0	7	1	UC24_WAIZE	P80630	zea mays (m
62	1	20.0	7	1	UF04_MOUSE	P38642	mus musculus
63	1	20.0	7	1	UH11_RAT	P58576	rattus norv
64	1	20.0	7	1	UN06_PINPS	P81675	pinus pinas
65	1	20.0	7	1	WMA1_ACHFU	P35919	achatina fu
66	1	20.0	7	1	WMA2_ACHFU	P35920	achatina fu
67	1	20.0	7	1	WMA3_ACHFU	P35921	achatina fu
68	1	20.0	8	1	ACL_THUAL	P18691	thunnus alb
69	1	20.0	8	1	AKHG_GRYBI	P14086	gryllus bim
70	1	20.0	8	1	AKH_LIBAU	P25418	libellula a
71	1	20.0	8	1	AKH_MELML	P25423	melolontha
72	1	20.0	8	1	AKH_TABAT	P14595	tabanus atr
73	1	20.0	8	1	AL12_CARMA	P81815	carcinus ma
74	1	20.0	8	1	AL15_CARMA	P81818	carcinus ma
75	1	20.0	8	1	AL17_CARMA	P81820	carcinus ma
76	1	20.0	8	1	AL18_CARMA	P81821	carcinus ma
77	1	20.0	8	1	AL11_CVDPO	P82152	cydia pomon
78	1	20.0	8	1	ALL3_CVDPO	P82154	cydia pomon
79	1	20.0	8	1	ALL4_CVDPO	P41840	calliphora
80	1	20.0	8	1	ALL4_CVDPO	P82155	cydia pomon
81	1	20.0	8	1	ALL5_CVDPO	P41841	calliphora
82	1	20.0	8	1	ALL5_CVDPO	P82156	cydia pomon
83	1	20.0	8	1	ALL6_CVDPO	P82157	cydia pomon
84	1	20.0	8	1	ALL7_CARMA	P81809	carcinus ma
85	1	20.0	8	1	ALL8_CARMA	P81811	carcinus ma
86	1	20.0	8	1	CAD1_ENTFA	P13268	enterococcu
87	1	20.0	8	1	CCKN_MACEU	P30369	macropus eu
88	1	20.0	8	1	CLP_THICU	P80488	thiobacillu
89	1	20.0	8	1	COM2_CONPU	P58785	conus purpu
90	1	20.0	8	1	CPD1_ENTFA	P13269	enterococcu
91	1	20.0	8	1	FAR1_PANRE	P41872	panagrellus
92	1	20.0	8	1	FAR1_PENMO	P83316	penaeus mon
93	1	20.0	8	1	FAR3_HOMAM	P41486	homarus ane
94	1	20.0	8	1	FAR7_ASCSU	P43171	ascaris suu
95	1	20.0	8	1	FAR8_CALVO	P41863	calliphora
96	1	20.0	8	1	FUSS_FUSSO	P81010	fusarium so
97	1	20.0	8	1	GLUR_HUMAN	P02729	homo sapien
98	1	20.0	8	1	HTF1_PERAM	P04548	periplaneta
99	1	20.0	8	1	HTF_TENMO	P25419	tenebrio mo
100	1	20.0	8	1	LCK1_LEUMA	P21140	leucophaea
101	1	20.0	8	1	LCK2_LEUMA	P21141	leucophaea
102	1	20.0	8	1	LCK3_LEUMA	P21142	leucophaea
103	1	20.0	8	1	LCK4_LEUMA	P21143	leucophaea
104	1	20.0	8	1	LCK6_LEUMA	P19988	leucophaea
105	1	20.0	8	1	LCK7_LEUMA	P19989	leucophaea
106	1	20.0	8	1	LCK8_LEUMA	P19990	leucophaea

107 1 20.0 8 1 LMT2 LOCHI
 108 1 20.0 8 1 LPK LEUMA
 109 1 20.0 8 1 LPMS STAP
 110 1 20.0 8 1 ORNY ORCLI
 111 1 20.0 8 1 PKK2 PERAM
 112 1 20.0 8 1 RPCH PANBO
 113 1 20.0 8 1 RSI ERWCH
 114 1 20.0 8 1 RST MYCIT
 115 1 20.0 8 1 RT34 BOVIN
 116 1 20.0 8 1 UH09 RAT
 117 1 20.0 8 1 UPAA HUMAN
 118 1 20.0 8 1 UPAA HUMAN
 119 1 20.0 8 1 AL10 CARMA
 120 1 20.0 9 1 AL11 CARMA
 121 1 20.0 9 1 ALC CHRE
 122 1 20.0 9 1 BS43 SERPL
 123 1 20.0 9 1 CCAP CARMA
 124 1 20.0 9 1 CONO CONGE
 125 1 20.0 9 1 CONO CONST
 126 1 20.0 9 1 COW CONVE
 127 1 20.0 9 1 DNF1 LOCHI
 128 1 20.0 9 1 FAR2 CALVO
 129 1 20.0 9 1 FAR2 PANRE
 130 1 20.0 9 1 FAR3 CALVO
 131 1 20.0 9 1 FAR3 PENMO
 132 1 20.0 9 1 FAR4 PENMO
 133 1 20.0 9 1 FAR5 ASCSU
 134 1 20.0 9 1 FAR5 CALVO
 135 1 20.0 9 1 FAR5 PENMO
 136 1 20.0 9 1 FAR6 CALVO
 137 1 20.0 9 1 FAR7 CALVO
 138 1 20.0 9 1 FAR8 MACRS
 139 1 20.0 9 1 FAR9 ASCSU
 140 1 20.0 9 1 FARD CALVO
 141 1 20.0 9 1 FARP CALSI
 142 1 20.0 9 1 FIBB MACFU
 143 1 20.0 9 1 FIBB PAPAN
 144 1 20.0 9 1 FIBB PAPA
 145 1 20.0 9 1 FRF1 GARBU
 146 1 20.0 9 1 HUTU KLEAE
 147 1 20.0 9 1 ISOT CYPCA
 148 1 20.0 9 1 KNL3 BOMVA
 149 1 20.0 9 1 LITO LITAU
 150 1 20.0 9 1 LITR PHYRO

ALIGNMENTS

RESULT 1
 ID THYF_PIG STANDARD; PRT; 9 AA.
 AC P01355;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Thymic factor.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 NCBI_TaxID=9823;
 [1]
 RP SEQUENCE.
 RX MEDLINE=78026571; PubMed=914862;
 RC Pleau J.-M., Dardenne M., Blouquit Y., Bach J.-P.;
 RT "Structural study of circulating thymic factor: a peptide isolated
 RT from pig serum. II. Amino acid sequence.";
 RL J. Biol. Chem. 252:8045-8047(1977).
 CC -1- MISCELLANEOUS: THE BIOLOGICAL SOURCE(S) AND PHYSIOLOGICAL
 CC ACTIVITIES OF THIS PEPTIDE HAVE NOT BEEN DETERMINED.
 DR PIR; A01523; YPFG.
 KW Pyrrolidone carboxylic acid. PYRROLIDONE CARBOXYLIC ACID.
 MOD_RES 1 1

P22396 locusta mig
 P13049 leucophaea
 P23211 staphylococ
 P82455 oriconetes
 P82692 periplaneta
 P08939 pandalus bo
 P37985 erwinia chr
 P33564 mycobacteri
 P82929 bos taurus
 P56575 rattus norv
 P30087 homo sapien
 P30096 homo sapien
 P81813 carcinus ma
 P81814 carcinus ma
 P82678 chlamydomon
 P83375 serratia pl
 P38556 carcinus ma
 P05486 conus geogr
 P05487 conus stria
 P83047 conus ventr
 P16339 locusta mig
 P41857 calliphora
 P41873 panagrellus
 P41858 calliphora
 P83318 penaeus mon
 P83319 penaeus mon
 P43170 ascaris suu
 P41860 calliphora
 P83320 penaeus mon
 P41861 calliphora
 P41862 calliphora
 P83281 macrobrachi
 P43172 ascaris suu
 P41868 calliphora
 P38495 callinectes
 P19345 macaca fusc
 P19344 papio anubi
 P19343 papio hamad
 P83350 sarcophaga
 P12381 klebsiella
 P42993 cyprinus ca
 P83058 bombina var
 P08945 litoria aur
 P08946 phyllomedus

SQ SEQUENCE 9 AA; 876 MW; D500B87866C5B33D CRC64;
 Query Match 60.0%; Score 3; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 GGS 5
 DB 6 GGS 8
 RESULT 2
 ID EOSI_HUMAN STANDARD; PRT; 4 AA.
 AC P02731;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 21-JUL-1986 (Rel. 01, Last annotation update)
 DE Eosinophilotoxic peptides.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 [1]
 RP SEQUENCE.
 RX MEDLINE=76078412; PubMed=1060093;
 RA Goetzl E.J., Austen K.F.;
 RT "Purification and synthesis of eosinophilotoxic tetrapeptides of
 RT human lung tissue: identification as eosinophil chemotactic factor of
 RT anaphylaxis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 72:4123-4127(1975).
 CC -1- MISCELLANEOUS: THESE PEPTIDES ARE RELEASED FROM MAST CELLS IN LUNG
 CC (AND OTHER TISSUES) DURING HYPERSENSITIVITY REACTIONS
 CC (ANAPHYLAXIS). THEIR ACTIVITIES, PREFERENTIALLY AFFECTING
 CC EOSINOPHILS, INCLUDE CHEMOTAXIS, CHEMOTACTIC DEACTIVATION, RELEASE
 CC OF ENZYMES, AND STIMULATION OF THE HEXOSE MONOPHOSPHATE SHUNT.
 DR GO; GO:0030105; P:anaphylaxis; IDA.
 DR GO; GO:0006935; P:chemotaxis; IDA.
 FT VARIANT 1 1 V->A (IN OTHER PEPTIDE).
 FT /FTID=VAR_005201.
 SQ SEQUENCE 4 AA; 390 MW; 6B05B862A00000000 CRC64;
 Query Match 40.0%; Score 2; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 GS 5
 DB 2 GS 3
 RESULT 3
 ID OCP3_OCTMI STANDARD; PRT; 4 AA.
 AC P86649;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cardioactive peptides Ocp-3/Ocp-4.
 OS Octopus minor (Octopus).
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;
 OC Octopodiformes; Octopoda; Incirrata; Octopodidae; Octopus.
 NCBI_TaxID=89766;
 [1]
 RN SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.
 RC TISSUE=Brain;
 RX MEDLINE=20336815; PubMed=10876044;
 RA Iwakoshi E., Hisada M., Minakata H.;
 RT "Cardioactive peptides isolated from the brain of a Japanese octopus,
 RT Octopus minor.";
 RL Peptides 21:623-630(2000).
 CC -1- FUNCTION: Cardioactive; has both positive chronotropic and
 CC inotropic effects on the heart. Ocp-4 is a 1000 time less

CC active than Ocp-3.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- PM: Ocp-4 has D-Ser instead of L-Ser.
 CC -!- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI.
 KW Hormone; D-amino acid.
 FT MOD RES 2 D-SERINE (IN OCP-4).
 SQ SEQUENCE 4 AA; 463 MW; 6AB365B810000000 CRC64;

Query Match 40.0%; Score 2; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GS 5
 DB 1 GS 2

RESULT 4
 CIP1 MYTED STANDARD; PRT; 6 AA.
 AC P13736;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Contraction-inhibiting peptide I (MIP I).
 OS Mytilus edulis (Blue mussel).
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
 OC Mytiloidea; Mytilidae; Mytilus.
 OX NCBI_TaxID=6550;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Pedal ganglion;
 RX MEDLINE=88240357; PubMed=3377776;
 RA Hirata T., Kubota I., Iwasawa N., Takabatake I., Ikeda T., Muneoka Y.;
 RT "Structures and actions of Mytilus inhibitory peptides.";
 RL Biochem. Biophys. Res. Commun. 152:1376-1382(1988).
 CC -!- FUNCTION: INHIBITORY ACTION ON CONTRACTIONS IN SEVERAL MOLLUSCAN MUSCLES.
 CC -!- SIMILARITY: TO MIP II.
 DR PIR: A27696; A27696.
 KW Hormone; Amidation.
 FT MOD RES 6
 SQ SEQUENCE 6 AA; 637 MW; 72C9C68775B81000 CRC64;

Query Match 40.0%; Score 2; DB 1; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GS 5
 DB 1 GS 2

RESULT 5
 FAR5_HIRME STANDARD; PRT; 7 AA.
 AC P42564;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE FMRFamide-like neuropeptide GSKYMRP-amide.
 OS Hirudo medicinalis (Medicinal leech).
 OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
 OC Arynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
 OX NCBI_TaxID=6421;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=92195954; PubMed=1686933;
 RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
 RT "Identification of RFamide neuropeptides in the medicinal leech.";
 RL Peptides 12:897-908(1991).
 CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE) FAMILY.

KW Neuropeptide; Amidation.
 FT MOD RES 7
 SQ SEQUENCE 7 AA; 858 MW; 69D408B53387810 CRC64;
 AMIDATION.

Query Match 40.0%; Score 2; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GS 2
 DB 1 GS 2

RESULT 6
 LANC_CARUI STANDARD; PRT; 7 AA.
 AC P36960;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lantibiotic carnocin UI49 (Fragment).
 OS Carnobacterium sp. (strain UI49).
 OC Bacteria; Firmicutes; Lactobacillales; Carnobacteriaceae;
 OC Carnobacterium.
 OX NCBI_TaxID=35782;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=92321768; PubMed=1622206;
 RA Stoffels G., Nissen-Meyer J., Gudmundsdottir A., Sletten K., Holo H.,
 RA Nes I.F.;
 RT "Purification and characterization of a new bacteriocin isolated from a Carnobacterium sp.";
 RL Appl. Environ. Microbiol. 58:1417-1422(1992).
 CC -!- FUNCTION: LANTHONINE-CONTAINING PEPTIDE ANTIBIOTIC (LANTIBIOTIC).
 CC ACTIVE ON GRAM-POSITIVE BACTERIA.
 KW Antibiotic; Bacteriocin; Lantibiotic.
 FT NON_TER 7
 SQ SEQUENCE 7 AA; 786 MW; 741776D05B05B810 CRC64;

Query Match 40.0%; Score 2; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GS 5
 DB 1 GS 2

RESULT 7
 AL16_CARMA STANDARD; PRT; 8 AA.
 AC P81819;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Carcinustatin 16.
 OS Carcinus maenas (Common shore crab) (Green crab).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
 OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
 OX NCBI_TaxID=6759;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
 RX MEDLINE=98121193; PubMed=9461295;
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
 RA Thorpe A.;
 RT "Isolation and identification of multiple neuropeptides of the allatostatin superfamily in the shore crab Carcinus maenas.";
 RL Eur. J. Biochem. 250:727-734(1997).
 CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 KW Neuropeptide; Amidation; Multigene family.

FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 813 MW; 7C286B45AB76878 CRC64;
Query Match 40.0%; Score 2; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GG 2
DB 1 GG 2

RESULT 8
ALL9 CARMA STANDARD; PRT; 8 AA.
ID ALL9 CARMA
AC P81812;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 9.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubranchyura; Portunoidae; Portunidae; Carcinus.
OC NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RT TISSUE=Cerebral ganglion, and Thoracic ganglion;
RC MEDLINE=98121193; PubMed=9461295;
RX Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RA "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734 (1997).
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
CC FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 781 MW; 7C2879CDBC476878 CRC64;
Query Match 40.0%; Score 2; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GG 2
DB 1 GG 2

RESULT 9
LCK5 LEUMA STANDARD; PRT; 8 AA.
ID LCK5 LEUMA
AC F19987;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Leucokinin V (L-V).
OS Leucophaea maderae (Madeira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberioidea;
OC Blaberidae; Leucophaea.
OC NCBI_TaxID=6988;
RN [1]
RP SEQUENCE.
RT TISSUE=Head;
RC MEDLINE=87052651; PubMed=2877794;
RX Holman G.M., Cook B.J., Nachman R.J.;
RA "Isolation, primary structure, and synthesis of leucokinin V and VI:
RT myotropic peptides of Leucophaea maderae.";
RL Comp. Biochem. Physiol. 88C:27-30 (1987).
CC -!- FUNCTION: THIS CEPHALOMOTROPIC PEPTIDE STIMULATES CONTRACTILE
CC ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).
CC -!- SIMILARITY: TO THE OTHER LEUCOKININS.

DR PIR; JS0315; JS0315.
KW Neuropeptide; Amidation.
FT MOD_RES 8 8
SQ SEQUENCE 8 AA; 784 MW; 736365A5B9C865B8 CRC64;
Query Match 40.0%; Score 2; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 GS 5
DB 1 GS 2

RESULT 10
UF06 MOUSE STANDARD; PRT; 8 AA.
ID UF06 MOUSE
AC P38644;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Unknown protein from 2D-page of fibroblasts (p50) (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RT TISSUE=Fibroblast;
RC MEDLINE=95009907; PubMed=7523108;
RX Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
RA "Separation and sequencing of familial and novel murine proteins
RT using preparative two-dimensional gel electrophoresis.";
RL Electrophoresis 15:735-745 (1994).
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 5.2, ITS MW IS: 50 kDa.
CC NON TER 8
SQ SEQUENCE 8 AA; 817 MW; A35DD878676B05B1 CRC64;
Query Match 40.0%; Score 2; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GG 2
DB 5 GG 6

RESULT 11
VGLG HSV2B STANDARD; PRT; 8 AA.
ID VGLG HSV2B
AC P81780;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Glycoprotein G (fragment).
OS Herpes simplex virus (type 2 / strain B4327UR).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OC NCBI_TaxID=103921;
RN [1]
RP SEQUENCE.
RT Lijerqvist J.-A., Svennerholm B., Bergstrom T.;
RL Submitted (APR-1999) to the SWISS-PROT data bank.
CC -!- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1 AND
CC 2: GH, GB, GC, GD, GI, AND GE.
CC -!- MISCELLANEOUS: GLYCOPROTEIN G IS MUCH LARGER IN HSV-2 THAN IN
CC HSV-1.
KW Glycoprotein.
FT NON TER 8
SQ SEQUENCE 8 AA; 683 MW; 7B47686772C865B8 CRC64;
Query Match 40.0%; Score 2; DB 1; Length 8;

Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
Db 1 GS 2

RESULT 12

WPI_PERAT ID_WPI_PERAT STANDARD; PRT; 8 AA.
AC P83195;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Wall protein-1 (Pwp-1) (Fragment).
OS Perkinsus atlanticus.
OC Eukaryota; Alveolata; Perkinsa; Perkinsida; Perkinsidae; Perkinsus.
OX NCBI_TaxID=106964;
RN [1]
RP SEQUENCE, FUNCTION, SUBCELLULAR LOCATION, AND DEVELOPMENTAL STAGE.
RX MEDLINE=22044350; PubMed=12049410;
RA Montes J.F., Dufort M., Llado A., Garcia-Valero J.;
RT "Characterization and immunolocalization of a main proteinaceous component of the cell wall of the protozoan parasite Perkinsus atlanticus.";
RL Parasitology 124:477-484(2002).
CC -!- FUNCTION: Is a major protein component of the cell wall. May play a key role in the organization of the cell wall and in promoting the survival of this parasite.
CC -!- SUBCELLULAR LOCATION: Cell wall. Disulfide-linked to other cell wall components.
CC -!- DEVELOPMENTAL STAGE: Expressed throughout all walled developmental stages.
KW Cell wall.
FT NON_TER
SQ SEQUENCE 8 AA; 765 MW; F1787DD87B1AAB16 CRC64;

Query Match 40.0%; Score 2; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
Db 7 GG 8

RESULT 13

DI_NEPNO ID_DI_NEPNO STANDARD; PRT; 9 AA.
AC P24816;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-MAR-1992 (Rel. 21, Last annotation update)
DE Gastrin/cholecystokinin-like peptide DI.
OS Nephrops norvegicus (Norway lobster).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea; Nephropoidea; Nephropidae; Nephrops.
OX NCBI_TaxID=6829;
RN [1]
RP SEQUENCE.
RC TISSUE=Stomach;
RX MEDLINE=92082847; PubMed=1747388;
RA Favrel P., Kegel G., Sedlmeier D., Keller R., van Wormhoudt A.;
RT "Structure and biological activity of crustacean gastrointestinal peptides identified with antibodies to gastrin/cholecystokinin.";
RL Biochimie 73:1233-1239(1991).
CC -!- FUNCTION: MAY CONTROL DIGESTION PROCESSES IN CRUSTACEANS.
CC -!- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.
DR PIR; S47432; S47432.
KW Hormone.
SQ SEQUENCE 9 AA; 1038 MW; 60EC79CAB6D8787B CRC64;

Query Match 40.0%; Score 2; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
Db 3 GG 4

RESULT 14

DSIP_RABIT ID_DSIP_RABIT STANDARD; PRT; 9 AA.
AC P01158;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Delta sleep-inducing peptide (DSIP).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE.
RX MEDLINE=77185324; PubMed=862769;
RA Monnier M., Dudler L., Gachter R., Maier P.F., Tobler H.J., Schoenenberger G.A.;
RT "The delta sleep inducing peptide (DSIP). Comparative properties of the original and synthetic nonapeptide.";
RL Experientia 33:548-552(1977).
RN [2]
RP SEQUENCE, AND SYNTHESIS.
RX MEDLINE=79054421; PubMed=568769;
RA Schoenenberger G.A., Maier P.F., Tobler H.J., Wilson K., Monnier M.;
RT "The delta EEG (sleep)-inducing peptide (DSIP). XI. Amino-acid analysis, sequence, synthesis and activity of the nonapeptide.";
RL Pflugers Arch. 376:119-129(1978).
RN [3]
RP REVIEW.
RX MEDLINE=87175129; PubMed=3550726;
RA Graf M.V., Kastin A.J.;
RT "Delta-sleep-inducing peptide (DSIP): an update.";
RL Peptides 7:1165-1187(1986).
CC -!- FUNCTION: WHEN INFUSED INTO THE MESODIENCEPHALIC VENTRICLE OF RECIPIENT RABBITS INDUCES SPINDLE AND DELTA EEG ACTIVITY AND REDUCED MOTOR ACTIVITIES.
CC -!- MISCELLANEOUS: THIS PEPTIDE WAS OBTAINED FROM DIALYSATES OF OCCIPITAL VENOUS SINUS BLOOD FROM RABBITS KEPT ASLEEP BY ELECTRIC STIMULATION OF THE THALAMUS.
CC -!- DATABASE: NAME=Protein Spotlight;
CC NOTE=Issue 8 of March 2001;
CC WWW="http://www.expasy.org/spotlight/articles/sptlt008.html".
DR PIR; A01422; ODRB.
SQ SEQUENCE 9 AA; 849 MW; DDD365BDDAA8787D CRC64;

Query Match 40.0%; Score 2; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
Db 3 GG 4

RESULT 15

PAR6_MACRS ID_PAR6_MACRS STANDARD; PRT; 9 AA.
AC P83279;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE PMFamide-like neuropeptide FLP6 (DGRNPLRF-amide).
OS Macrobrachium rosenbergii (Giant fresh water prawn).

OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Caridea;
 OC Palaemonoidea; Palaemonidae; Macrobrachium.
 OX NCBI_TaxID=79674;

RN
 RN SEQUENCE, AND MASS SPECTROMETRY.

TISSUE=Eyestalk;
 RC MEDLINE=21107394; PubMed=11179812;
 RX Sthigorngul P., Sarathongkul W., Longyant S., Panchan N.,
 RA Sthigorngul W., Petson A.;

RT "Three more novel FMRP-like neuropeptide sequences from the
 RT eyestalk of the giant freshwater prawn Macrobrachium rosenbergii.";
 RL Peptides 22:191-197(2001).
 CC -!- MASS SPECTROMETRY: MW=1080.7; METHOD=MALDI.

CC -!- SIMILARITY: BELONGS TO THE FARP (FMRP-RELATED PEPTIDE)
 CC FAMILY.

GO: GO:0007218; P-neuropeptide signaling pathway; IDA.

KW Neuropeptide; Amidation.

FT MOD RES 9 9 AMIDATION.

SQ SEQUENCE 9 AA; 1081 MW; 26800729C4540878 CRC64;

Query Match 40.0%; Score 2; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GG 2
 DB 2 GG 3

RESULT 16

ID FIBB_ERYPA STANDARD; PRT; 9 AA.

AC P19346;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).

GN FGB.

OC Erythrocytes patas (Red guenon) (Huesar).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;

OC Cercopithecoidea; Erythrocytes.

FT NON_TER 9 9

SQ NCBI_TaxID=9538;

[1]

Query Match 40.0%; Score 2; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GG 2
 DB 2 GG 3

RESULT 17

ID FIBB_THEGE STANDARD; PRT; 9 AA.

AC P19342;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).

GN FGB.

OC Theropithecus gelada (Gelada baboon).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;

OC Cercopithecoidea; Theropithecus.

FT NON_TER 9 9

SQ NCBI_TaxID=9565;

[1]

Query Match 40.0%; Score 2; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GG 2
 DB 7 GG 8

Query Match 40.0%; Score 2; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GG 2
 DB 7 GG 8

RESULT 18

ID NEF_HV128 STANDARD; PRT; 9 AA.

AC P12481;

DT 01-OCT-1989 (Rel. 12, Created)

DT 01-OCT-1989 (Rel. 12, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Negative factor (P-protein) (27 kDa protein) (3'ORF) (Fragment).

GN NEF.

OC Human immunodeficiency virus type 1 (Z-84 isolate) (HIV-1).

OC Viruses; Retroviridae; Lentivirus.

FT NON_TER 9 9

SQ NCBI_TaxID=11681;

[1]

Query Match 40.0%; Score 2; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GG 2
 DB 7 GG 8

RL AIDS Res. Hum. Retroviruses 4:165-173(1988).

CC -!- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating

CC activities. It seems to down-regulate the CD4(T4) antigen.

CC -!- MISCELLANEOUS: THE Z-84 ISOLATE WAS TAKEN FROM A 54 YEAR-OLD

CC ZAIKIAN MALE.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; J03653; AAA44687.1; -.

DR HIV; J03653; NEFSJY1

KW AIDS; Myristate; GTP-binding.

FT LIPID 2 2 MYRISTATE (BY SIMILARITY).

FT NON_TER 9 9

SQ SEQUENCE 9 AA; 967 MW; 319CB325A3733878 CRC64;

Query Match 40.0%; Score 2; DB 1; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.3e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2

DB 2 GG 3

RESULT 19

XYLA_STRSQ STANDARD; PRT; 9 AA.

AC P19149;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Xylose isomerase (EC 5.3.1.5) (Fragment).

GN XYLA.

OS Streptomyces sp. (strain NCL 82-5-1).

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Streptomycineae; Streptomycetaceae; Streptomycetes.

OX NCBI_TaxID=1931;

RN [1]

RP SEQUENCE.

RX MEDLINE=88326335; PubMed=3415697;

RA Pawar H.S., Kannan K., Srinivasan M.C., Vartak H.G.;

RT "Purification and characterisation of glucose (xylose) isomerase from

RT Chainia sp. (NCL 82-5-1).";

RL Biochem. Biophys. Res. Commun. 155:411-417(1988).

CC -!- FUNCTION: Involved in D-xylose catabolism.

CC -!- CATALYTIC ACTIVITY: D-xylose = D-xylulose.

CC -!- COFACTOR: Binds 2 magnesium ions per subunit (Potential).

CC -!- SUBUNIT: Homotetramer.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC -!- SIMILARITY: BELONGS TO THE XYLOSE ISOMERASE FAMILY.

DR PIR; A31576; A31576.

DR HAMAP; MF_00455; -; 1.

DR InterPro; IPR001998; Xylose isom.

DR PROSITE; PS00172; XYLOSE ISOMERASE_1; PARTIAL.

DR PROSITE; PS00173; XYLOSE ISOMERASE_2; PARTIAL.

KW Isomerase; Pentose shunt; Xylose metabolism; Metal-binding; Magnesium.

FT NON_TER 9 9

SQ SEQUENCE 9 AA; 983 MW; F64BA1EDC5B87DD1 CRC64;

Query Match 40.0%; Score 2; DB 1; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.3e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5

DB 4 GS 5

RESULT 20

AMPN_HELAM STANDARD; PRT; 10 AA.

AC P81731;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Amino-peptidase N (EC 3.4.11.2) (CRYIA(C) receptor) (Fragment).

OS Helicoverpa armigera (Cotton bollworm).

CC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;

CC Neoptera; Endopterygota; Lepidoptera; Glossata; Noctuoidea;

CC Noctuidae; Heliothinae; Helicoverpa.

OX NCBI_TaxID=29058;

RN [1]

RP SEQUENCE.

RC TISSUE=Midgut;

RA Ingle S.S., Trivedi N., Praaad R., Rao K.K., Chatpar H.S.;

RT "Amino-peptidase-N as a receptor for Bacillus thuringiensis CryIAC

RT toxin from Helicoverpa armigera.";

RL Submitted (MAR-1999) to the SWISS-PROT data bank.

CC -!- FUNCTION: ACTS AS A RECEPTOR FOR B.THURINGIENSIS TOXIN CRYIAC.

CC -!- CATALYTIC ACTIVITY: Release of an N-terminal amino acid, Xaa-|-

CC Xbb-from a peptide, amide or arylamide. Xaa is preferably Ala, but

CC may be most amino acids including Pro (slow action). When a

CC terminal hydrophobic residue is followed by a prolyl residue, the

CC two may be released as an intact Xaa-Pro dipeptide.

CC -!- COFACTOR: Binds 1 zinc ion (By similarity).

CC -!- SIMILARITY: Belongs to peptidase family M1.

DR InterPro; IPR006025; Zn_MTPeptidse.

DR PROSITE; PS00142; ZINC_PROTEASE; PARTIAL.

KW Hydrolase; Metalloprotease; Amino-peptidase; Zinc.

FT NON_TER 10 10

SQ SEQUENCE 10 AA; 1093 MW; 05042EB87B11F1B8 CRC64;

Query Match 40.0%; Score 2; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 5.8e+03;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5

DB 7 GS 8

RESULT 21

COXH_ONCMY STANDARD; PRT; 10 AA.

AC P80331;

DT 01-OCT-1994 (Rel. 30, Created)

DT 01-OCT-1994 (Rel. 30, Last sequence update)

DT 01-OCT-1994 (Rel. 30, Last annotation update)

DE Cytochrome c oxidase polypeptide Vic (EC 1.9.3.1) (Fragment).

OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Actinopterygii; Neopterygii; Teleostei; Euteleostei;

CC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.

OX NCBI_TaxID=8022;

RN [1]

RP SEQUENCE.

RC TISSUE=Liver;

RX MEDLINE=94237150; PubMed=8181469;

RA Freund R., Kadenbach B.;

RT "Identification of tissue-specific isoforms for subunits Vb and VIa

RT of cytochrome c oxidase isolated from rainbow trout.";

RL Eur. J. Biochem. 221:1111-1116(1994).

CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE

CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN

CC MITOCHONDRIAL ELECTRON TRANSPORT.

CC -!- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) = 4 ferrocyclochrome

CC c + 2 H(2)O.

CC -!- SUBCELLULAR LOCATION: Mitochondrial inner membrane.

DR PIR; S43630; S43630.

KW Oxidoreductase; Inner membrane; Mitochondrion.

```

T  NON TER      10      10
SQ SEQUENCE 10 AA; 977 MW; E11B4076DC772DA CRC64;

Query Match      40.0%; Score 2; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 5.8e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
   ||
DB 8 GG 9

RESULT 22
CU30-LOCMI STANDARD; PRT; 10 AA.
AC P11735;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Cuticle protein 30 (LM-30) (LM-ACP 30) (Fragment).
DC Locusta migratoria (Migratory locust).
CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
CC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;
CC Acridoidea; Acrididae; Oedipodinae; Locusta.
CX NCBI_TaxID=7004;
RN [1]
RP SEQUENCE.
RR MEDLINE=86108304; PubMed=3943519;
RA Hoejrup P., Andersen S.O., Roepstorff P.;
RT "Isolation, characterization, and N-terminal sequence studies of
RT cuticular proteins from the migratory locust, Locusta migratoria.";
RL Eur. J. Biochem. 154:153-159(1986).
CC -!- FUNCTION: COMPONENT OF THE CUTICLE OF MIGRATORY LOCUST WHICH
CC CONTAINS MORE THAN 100 DIFFERENT STRUCTURAL PROTEINS.
CC PIR; H24802; H24802.
DR Structural protein; Cuticle.
KW NON TER 10 10
SQ SEQUENCE 10 AA; 969 MW; 4973E36B58772877 CRC64;

Query Match      40.0%; Score 2; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 5.8e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y 1 GG 2
   ||
DB 8 GG 9

RESULT 23
ESTA SCHGA STANDARD; PRT; 10 AA.
AC P81012;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Esterase 52 kDa subunit (EC 3.1.1.1) (Carboxylic-ester hydrolase)
DE (Fragment).
DE DE
OS Schizaphis graminum (Aphid).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
OC Aphidoidea; Aphididae; Aphidini; Schizaphis.
CX NCBI_TaxID=13262;
RN [1]
RP SEQUENCE.
RX MEDLINE=97468499; PubMed=9327586;
RA Siegfried B.D., Ono M., Swanson J.J.;
RT "Purification and characterization of a carboxylesterase associated
RT with organophosphate resistance in the greenbug, Schizaphis graminum
RT (Homoptera: Aphididae).";
RL Arch. Insect Biochem. Physiol. 36:229-240(1997).
CC -!- CATALYTIC ACTIVITY: A carboxylic ester + H(2)O = an alcohol + a
CC carboxylic anion.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

```

```

CC -!- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
DR InterPro; IPR002018; CarbesteraseB.
DR PROSITE; PS00122; CARBOXYLESTERASE_B_1; PARTIAL.
DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; PARTIAL.
KW Hydrolase; Serine esterase.
FT NON TER 10 10
SQ SEQUENCE 10 AA; 1025 MW; 018ABE587865A2C0 CRC64;

Query Match      40.0%; Score 2; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 5.8e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
   ||
DB 6 GG 7

RESULT 24
GRP_RANR1 STANDARD; PRT; 10 AA.
AC P23260;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Neuromedin C.
OS Rana ridibunda (Laughing frog) (Marsh frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
CX NCBI_TaxID=8406;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=91315477; PubMed=1859413;
RA Conlon J.M., O'Harte F., Vaudry H.;
RT "Primary structures of the bombesin-like neuropeptides in frog brain
RT show that bombesin is not the amphibian gastrin-releasing peptide.";
RL Biochem. Biophys. Res. Commun. 178:526-530(1991).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE BOMBESIN/NEUROMEDIN B/RANATENSIN
CC FAMILY.
DR PIR; PQ0177; PQ0177.
DR InterPro; IPR000874; Bombesin.
DR Pfam; PF02044; Bombesin; 1.
DR PROSITE; PS00257; BOMBESIN; 1.
KW Bombesin family; Amidation.
FT MOD RES 10 10
SQ SEQUENCE 10 AA; 1094 MW; F91FBAE862CDC371 CRC64;

Query Match      40.0%; Score 2; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 5.8e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
   ||
DB 1 GS 2

RESULT 25
NO40-TOBAC STANDARD; PRT; 10 AA.
AC P55962;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Early nodulin 40 homolog.
GN ENOD40.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; Lamiales; Solanales; Solanaceae; Nicotiana.
CX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=cv. Petit Havana SRI;
RX MEDLINE=96320417; PubMed=86662527;
RA van de Sande K., Pawlowski K., Czaja I., Wienieke U., Schell J.,
RA Schmidt J., Walden R., Matvienko M., Wellink J., van Kammen A.,
RA Fransen H., Bisseling T.;
RT "Modification of phytohormone response by a peptide encoded by ENOD40
RT of legumes and a nonlegume.";
RL Science 273:370-373(1996).
CC -!- FUNCTION: MODULATES THE ACTION OF AUXIN AND MAY FUNCTION AS PLANT
CC GROWTH REGULATOR THAT ALTERS PHYTOHORMONE RESPONSES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X98716; CAA67267.1; -.
KW Modulation.
SQ SEQUENCE 10 AA; 1173 MW; 2AB248E05DDB1AB3 CRC64;

Query Match 40.0%; Score 2; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 5.8e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
Db 9 GS 10

RESULT 26
RT02 BOVIN STANDARD; PRT; 10 AA.
AC P82923;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Mitochondrial 28S ribosomal protein S2 (MRP-S2) (Fragment).
GN MRPS2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Theria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=21276436; PubMed=11279123;
RA Koc E.C., Butkhat W., Blackburn K., Moseley A., Spremulli L.L.;
RT "The small subunit of the mammalian mitochondrial ribosome:
RT identification of the full complement of ribosomal proteins present.";
RL J. Biol. Chem. 276:19363-19374(2001)
CC -!- SUBUNIT: Component of the mitochondrial ribosome small subunit
CC (28S) which comprises a 12S rRNA and about 30 distinct proteins.
CC -!- SUBCELLULAR LOCATION: Mitochondrial.
CC -!- SIMILARITY: BELONGS TO THE S2P FAMILY OF RIBOSOMAL PROTEINS.
DR InterPro; IPR001865; Ribosomal S2.
DR Pfam; PF00318; Ribosomal S2; PARTIAL.
DR PROSITE; PS00962; RIBOSOMAL S2.1; PARTIAL.
KW Ribosomal protein; Mitochondrion.
FT NON_TER 1
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 1246 MW; 6A7A6673C04B476B CRC64;

Query Match 40.0%; Score 2; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 5.8e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
Db 8 GS 9

RESULT 27
TK01_UREUN STANDARD; PRT; 10 AA.
AC P40751;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Urechis tachykinin I.
OS Urechis unicinctus.
OC Eukaryota; Metazoa; Echiura; Xenopneusta; Urechidae; Urechis.
OX NCBI_TaxID=6432;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC TISSUE=ventral nerve cord;
RX MEDLINE=93236558; PubMed=8476410;
RA Ikeda T., Minakata H., Nomoto K., Kubota I., Muneoka Y.;
RT "Two novel tachykinin-related neuropeptides in the echiuroid worm,
RT Urechis unicinctus.";
RL Biochem. Biophys. Res. Commun. 192:1-6(1993).
CC -!- FUNCTION: CONTRACTILE ACTION ON THE INNER CIRCULAR BODY-WALL
CC MUSCLE OF THE ANIMAL.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: SOME SIMILARITY TO TACHYKININS.
KW Tachykinin; Neuropeptide; Amidation.
FT MOD_RES 10 10
FT SEQUENCE 10 AA; 1177 MW; C6D1C462C9D6C5A6 CRC64;

Query Match 40.0%; Score 2; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 5.8e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
Db 8 GS 9

RESULT 28
TPIS_NICPL STANDARD; PRT; 10 AA.
AC P19118;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Triosephosphate isomerase, cytosolic (EC 5.3.1.1) (TIM) (Fragment).
OS Nicotiana glauca (Leadwort-leaved tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; Lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4092;
RN [1]
RP SEQUENCE.
RA Bauw G., de Loose M., Inze D., van Montagu M., Vandekerckhove J.;
RT "Alterations in the phenotype of plant cells studied by NH2-terminal
RT amino acid-sequence analysis of proteins electrophoretically separated from two-
RT dimensional gel-separated total extracts.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:4806-4810(1987).
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate = glyceraldehyde
CC 3-phosphate.
CC -!- PATHWAY: Plays an important role in several metabolic pathways.
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- MISCELLANEOUS: IN PLANTS, THERE ARE TWO TYPES OF TPIS, CYTOSOLIC
CC AND PLASTID.
CC -!- SIMILARITY: BELONGS TO THE TRIOSEPHOSPHATE ISOMERASE FAMILY.
DR PIR; A27617; A27617.
DR InterPro; IPR000652; Triophos_ismrse.
DR PROSITE; PS00171; TIM; PARTIAL.
KW Isomerase; Glycolysis; Gluconeogenesis; Fatty acid biosynthesis;
KW Pentose shunt.
FT NON_TER 10 10
FT SEQUENCE 10 AA; 1140 MW; 80B9D37862C9C9D1 CRC64;

```

Query Match 40.0%; Score 2; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 5.8e+03; Indels 0; Gaps 0;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
 ||
 Db 7 GG 8

RESULT 29
 UP08 HUMAN STANDARD; PRT; 10 AA.
 ID AC UP08 HUMAN
 AC P30094;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Unknown protein from 2D-page of plasma (Spot 34) (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.

RC TISSUE=Plasma;
 RX MEDLINE=93092937; PubMed=1459097;
 RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,
 Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,
 Hochstrasser D.F.;
 RA "Plasma protein map: an update by microsequencing."
 RL Electrophoresis 13:707-714(1992).
 CC -!- MISCELLANEOUS; ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 PROTEIN IS: 7.2, ITS MW IS: 16 kDa.
 DR SWISS-2DPAGE; P30094; HUMAN.
 FT NON_TER 1 1
 FT VART 4 4 S -> H.
 FT NON_TER 10 10 /FTID=VAR_000003.
 FT SEQUENCE 10 AA; 977 MW; 2EA6E0C77AE325B8 CRC64;

Query Match 40.0%; Score 2; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 5.8e+03; Indels 0; Gaps 0;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
 ||
 Db 3 GS 4

RESULT 30
 GRMM HUMAN STANDARD; PRT; 3 AA.
 ID GRMM HUMAN
 AC P01157;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 21-JUL-1986 (Rel. 01, Last annotation update)
 DE Growth-modulating peptide.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.

RC MEDLINE=77162369; PubMed=858356;
 RX Schlesinger D.H., Pickart L., Thaler M.M.;
 RA "Growth-modulating serum tripeptide is glycyl-histidyl-lysine."
 RL Experientia 33:324-325(1977).
 CC -!- MISCELLANEOUS; THIS SERUM TRIPEPTIDE HAS BEEN FOUND TO STIMULATE
 GROWTH OF SOME CELL TYPES AND TO INHIBIT OTHER TYPES IN VITRO.
 DR GO:0001558; P:regulation of cell growth; NAS.
 SQ SEQUENCE 3 AA; 340 MW; 6331E81000000000 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 3;

Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 |
 Db 1 G 1

RESULT 31
 ACHI_ACHFU STANDARD; PRT; 4 AA.
 ID ACHI_ACHFU
 AC P35904;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Achatin-I.
 OS Achatina fulica (Giant African snail).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
 OC Sigmurethra; Achatinoidea; Achatinidae; Achatina.
 NCBI_TaxID=6530;
 RN [1]
 RP SEQUENCE, CHARACTERIZATION, AND SYNTHESIS.

RC STRAIN=Perussac; TISSUE=Ganglion;
 RX MEDLINE=89273551; PubMed=2597281;
 RA Kamatani Y., Minakata H., Kenny P.T.M., Iwashita T., Watanabe K.,
 Funase K., Sun X.P., Yongsiri A., Kim K.H., Novales-Li P.,
 Novales E.T., Kanapi C.G., Takeuchi H., Nomoto K.;
 RA "Achatin-I, an endogenous neuroexcitatory tetrapeptide from Achatina
 fulica Perussac containing a D-amino acid residue."
 RL Biochem. Biophys. Res. Commun. 160:1015-1020(1989).
 RN [2]
 RP CHARACTERIZATION.
 RC STRAIN=Perussac; TISSUE=Heart atrium;
 RX MEDLINE=91264856; PubMed=1675568;
 RA Fujimoto K., Kubota I., Yasuda-Kamatani Y., Minakata H., Nomoto K.,
 Yoshida M., Harada A., Muneoka Y., Kobayashi M.;
 RA "Purification of achatin-I from the atria of the African giant snail,
 Achatina fulica, and its possible function."
 RL Biochem. Biophys. Res. Commun. 177:847-853(1991).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY.

RX MEDLINE=93014529; PubMed=1399265;
 RA Ishida T., In Y., Doi M., Inoue M., Yasuda-Kamatani Y., Minakata H.,
 Iwashita T., Nomoto K.;
 RA "Crystal structure and molecular conformation of achatin-I
 (H-Gly-D-Phe-Ala-Asp-OH), an endogenous neuropeptide containing a
 D-amino acid residue."
 RL Int. J. Pept. Protein Res. 39:258-264(1992).
 CC -!- FUNCTION: NEUROEXCITATORY PEPTIDE; INCREASES THE IMPULSE FREQUENCY
 AND PRODUCES A SPIKE BROADENING OF THE IDENTIFIED HEART EXCITATORY
 NEURON (PON); ALSO ENHANCES THE AMPLITUDE AND FREQUENCY OF THE
 HEART BEAT. HAS ALSO AN EFFECT ON SEVERAL OTHER MUSCLES.
 DR PIR: A32480; A32480.
 KW Hormone; D-amino acid.
 FT MOD_RES 2 2 D-PHENYLALANINE.
 SQ SEQUENCE 4 AA; 408 MW; 6AADD9C810000000 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05; Indels 0; Gaps 0;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 |
 Db 1 G 1

RESULT 32
 DCML_PSECH STANDARD; PRT; 4 AA.
 ID DCML_PSECH
 AC P19916;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Carbon monoxide dehydrogenase large chain (EC 1.2.99.2) (CO
DE dehydrogenase subunit L) (CO-DH L) (Fragment).
GNTL.

OS Pseudomonas carboxydohydrogena.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Bradyrhizobiaceae.

OX NCBI_TaxID=290;

RN [1]

RP SEQUENCE.

RX MEDLINE=90055678; PubMed=2818128;

RA Kraut M., Hugendieck I., Herwig S., Meyer O.;

RT "Homology and distribution of CO dehydrogenase structural genes in

RT carboxydohydrogenic bacteria.";

RL Arch. Microbiol. 152:335-341(1989).

CC -!- FUNCTION: Catalyzes the oxidation of carbon monoxide to carbon

CC dioxide.

CC -!- CATALYTIC ACTIVITY: CO + H(2)O + acceptor = CO(2) + reduced

CC acceptor.

CC -!- COFACTOR: Molybdenum (molybdopterin).

CC -!- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND

CC SMALL.

DR PIR: PLO140; PLO140.

KW Oxidoreductase; Molybdenum.

FT NON_TER 4 4

SQ SEQUENCE 4 AA; 441 MW; 77618876F0000000 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 4;

Best Local Similarity 100.0%; Pred.No. 1.3e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 G 1

Db 2 G 2

RESULT 33

OCPI_OCTMI

ID OCPI_OCTMI STANDARD; PRT; 4 AA.

AC P58648;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Cardioactive peptides Ocp-1/Ocp-2.

OS Octopus minor (Octopus).

OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;

OC Octopodiformes; Octopoda; Incirrata; Octopodidae; Octopus.

OX NCBI_TaxID=89766;

RN [1]

RP SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.

RX TISSUE=Brain;

RX MEDLINE=20336815; PubMed=10876044;

RA Iwakoshi E., Hisada M., Minakata H.;

RA "Cardioactive peptides isolated from the brain of a Japanese octopus,

RT Octopus minor.";

RL Peptides 21:623-630(2000).

CC -!- FUNCTION: Cardioactive; has both positive chronotropic and

CC inotropic effects on the heart. Ocp-2 is a 1000 time less

CC active than Ocp-1.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- PTM: Ocp-2 has 1-Phe instead of D-Phe.

CC -!- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI.

KW Hormone; D-amino acid.

FT MOD_RES 2 2

SQ SEQUENCE 4 AA; 394 MW; 6A879C8100000000 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 4;

Best Local Similarity 100.0%; Pred.No. 1.3e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 G 1

Db 1 G 1

RESULT 34

RM01_YEAST

ID RM01_YEAST STANDARD; PRT; 4 AA.

AC P36515;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 01-JUN-1994 (Rel. 29, Last annotation update)

DE Mitochondrial 60S ribosomal protein L1 (Yml1) (Fragment).

GN MRPL1.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI_TaxID=4932;

RN [1]

RP SEQUENCE.

RX MEDLINE=91285106; PubMed=2060626;

RA Grohmann L., Graack H.-R., Kruff V., Choli T., Goldschmidt-Reisin S.;

RA Kitakawa M.;

RT "Extended N-terminal sequencing of proteins of the large ribosomal

RT subunit from yeast mitochondria.";

RL FEBS Lett. 284:51-56(1991).

DR PIR: S17255; S17255.

DR SGB: L0002681; MRPL1.

KW Ribosomal protein; Mitochondrion.

FT NON_TER 4 4

SQ SEQUENCE 4 AA; 402 MW; 7771B2D5D0000000 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 4;

Best Local Similarity 100.0%; Pred.No. 1.3e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 S 5

Db 1 S 1

RESULT 35

AL14_CARMA

ID AL14_CARMA STANDARD; PRT; 5 AA.

AC P81817;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Carcinustatin 14.

OS Carcinus maenas (Common shore crab) (Green crab).

OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;

OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;

OC Eubrachyura; Portunioidea; Portunidae; Carcinus.

OX NCBI_TaxID=6759;

RN [1]

RP SEQUENCE.

RC TISSUE=Cerebral ganglion, and Thoracic ganglion;

RX MEDLINE=98121193; PubMed=9461295;

RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.;

RA "Isolation and identification of multiple neuropeptides of the

RT allatostatin superfamily in the shore crab Carcinus maenas.";

RL Eur. J. Biochem. 250:727-734(1997).

CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.

CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.

KW Neuropeptide; Amidation; Multigene family.

FT MOD_RES 5 5

SQ SEQUENCE 5 AA; 586 MW; 672879D5AB300000 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 5;

Best Local Similarity 100.0%; Pred.No. 1.3e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 S 5

Db 2 S 2


```

RESULT 36
BIOB_CITR STANDARD; PRT; 5 AA.
ID BIOB_CITR STANDARD; PRT; 5 AA.
AC P12997;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Biotin synthase (EC 2.8.1.6) (Biotin synthetase) (Fragment).
GN BIOD.
OS Citrobacter freundii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Citrobacter.
OX NCBI_TaxID=546;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=89006280; PubMed=2971595;
SHUAN D., Campbell A.;
"Transcriptional regulation and gene arrangement of Escherichia coli,
Citrobacter freundii and Salmonella typhimurium biotin operons.";
Gene 67:203-211(1988).
CC -I- CATALYTIC ACTIVITY: Dethiobiotin + sulfur = biotin.
CC -I- PATHWAY: Biotin biosynthesis; last step.
CC -I- SIMILARITY: BELONGS TO THE BIOTIN AND LIPOIC ACID SYNTHETASES
FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC -----
CC ENBL; M21922; -; NOT_ANNOTATED_CDS.
CC PIR; I40698.
CC Biotin biosynthesis; Iron-sulfur; Transferase.
CC NON TER 5
CC SEQUENCE 5 AA; 532 MW; 75A591EDD6F00000 CRC64;
CC
Query Match 20.0%; Score 1; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 S 5
DB 4 S 4

```

```

RESULT 37
PAP2_PAPMA STANDARD; PRT; 5 AA.
ID PAP2_PAPMA STANDARD; PRT; 5 AA.
AC P81854;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Pardaxin II (PXII) (Fragment).
OS Pardaxinus marmoratus (Red sea moses sole).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Pleuronectiformes;
OC Soleioidae; Soleidae; Pardaxinus.
OX NCBI_TaxID=31087;
RN [1]
SEQUENCE.
CC TISSUE=Skin secretion;
CC MEDLINE=87057369; PubMed=3782138;
CC Lazarovici P., Primor N., Loew L.M.;
CC "Purification and pore-forming activity of two hydrophobic
CC polypeptides from the secretion of the Red sea moses sole (Pardaxinus
CC marmoratus).";
CC J. Biol. Chem. 261:16704-16713(1986).
CC -I- FUNCTION: Exhibits unusual shark repellent and surfactant

```

```

CC properties. Forms voltage-dependent, ion-permeable channels
CC in membranes. At high concentration causes cell membrane lysis.
CC -I- SUBUNIT: Monomer. In aqueous solution exists as a tetramer.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- SIMILARITY: BELONGS TO THE PARDAXIN FAMILY.
KW Toxin.
FT NON TER 5
SQ SEQUENCE 5 AA; 614 MW; 7769C9C8100000 CRC64;
Query Match 20.0%; Score 1; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 G 1
DB 1 G 1

```

```

RESULT 38
RE32_LITRU STANDARD; PRT; 5 AA.
ID RE32_LITRU STANDARD; PRT; 5 AA.
AC P82073;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Rubellidin 3.2.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Pelodyadinae; Litoria.
OX NCBI_TaxID=104895;
RN [1]
SEQUENCE.
CC TISSUE=Skin secretion;
CC Wamnitz F.A., Bowie J.H., Tyler M.J., Wallace J.C.;
CC "Peptides from the skin glands of the Australian buzzing tree frog
CC Litoria electrica. Comparison with the skin peptides from Litoria
CC rubella.";
CC Aust. J. Chem. 52:639-645(1999).
CC -I- FUNCTION: Shows neither neuropeptide activity nor antibiotic
CC activity.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.
KW Amphibian defense peptide.
SQ SEQUENCE 5 AA; 570 MW; 71A9C9C862A00000 CRC64;
Query Match 20.0%; Score 1; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 G 1
DB 2 G 2

```

```

RESULT 39
TPIS_CANFA STANDARD; PRT; 5 AA.
ID TPIS_CANFA STANDARD; PRT; 5 AA.
AC P54714;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Triosephosphate isomerase (EC 5.3.1.1) (TIM) (Fragment).
GN TP11.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
SEQUENCE.
CC TISSUE=Heart;
CC MEDLINE=98163340; PubMed=9504812;
CC Dunn M.J., Corbett J.M., Wheeler C.H.;

```

RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
 RT dog heart proteins.";
 RL Electrophoresis 18:2795-2802(1997).
 CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate = glyceralone
 CC phosphate.
 CC -!- PATHWAY: Plays an important role in several metabolic pathways.
 CC -!- SUBUNIT: Homodimer (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE TRIOSEPHOSPHATE ISOMERASE FAMILY.
 DR HSC-2DPAGE; P54714; DOG.
 DR InterPro; IPR000652; triophos_ismrse.
 DR PROSITE; PS00171; TIM; PARTIAL.
 KW Isomerase; Glycolysis; Gluconeogenesis; Fatty acid biosynthesis;
 KW Pentose shunt.
 FT NON_TER 1 1
 FT NON_TER 5 5
 SQ SEQUENCE 5 AA; 550 MW; 64444862C9A00000 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 G 1
 Db 3 G 3

RESULT 40

UF01_MOUSE
 ID UF01_MOUSE STANDARD; PRT; 5 AA.
 AC P38639;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE Unknown protein from 2D-page of fibroblasts (P19) (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Fibroblast; PubMed=7523108;
 RX MEDLINE=95009907; PubMed=7523108;
 RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
 RT "Separation and sequencing of familial and novel murine proteins
 RT using preparative two-dimensional gel electrophoresis.";
 RL Electrophoresis 15:735-745(1994).
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 6.6, ITS MW IS: 19 KDa.
 FT NON_TER 5 5
 SQ SEQUENCE 5 AA; 717 MW; 7364087043100000 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 G 1
 Db 3 G 3

RESULT 41

UXA4_CHLTR
 ID UXA4_CHLTR STANDARD; PRT; 5 AA.
 AC P38005;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Unknown protein from 2D-page from elementary body (Fragment).
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=813;
 RN [1]
 RP SEQUENCE.

RC STRAIN=L2/434/Bu;
 RA Bini L., Santucci A., Magi B., Marzocchi B., Sanchez-Campillo M.,
 RA Comanducci M., Christianen G., Birkeund S., Viretton S., Ratti G.,
 RA Pallini V.;
 RL Submitted (SEP-1994) to the SWISS-PROT data bank.
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 4.5, ITS MW IS: 28 kDa.
 DR Siena-2DPAGE; P38005; -.
 FT NON_TER 5 5
 SQ SEQUENCE 5 AA; 474 MW; 758AA865AA800000 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 5 S 5
 Db 2 S 2

RESULT 42

CIP2_MYTED
 ID CIP2_MYTED STANDARD; PRT; 6 AA.
 AC P13737;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Contraction-inhibiting peptide II (MIP II).
 OS Mytilus edulis (Blue mussel).
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
 OC Mytiloidea; Mytilidae; Mytilus.
 OX NCBI_TaxID=6550;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Pedal ganglion;
 RX MEDLINE=8240357; PubMed=3377776;
 RA Hirata T., Kubota I., Iwasawa N., Takabatake I., Ikeda T., Muneoka Y.;
 RT "Structures and actions of Mytilus inhibitory peptides.";
 RL Biochem. Biophys. Res. Commun. 152:1376-1382(1988).
 CC -!- FUNCTION: INHIBITORY ACTION ON CONTRACTIONS IN SEVERAL MOLLUSCAN
 CC MUSCLES.
 CC -!- SIMILARITY: TO MIP I.
 DR PIR; B27696; B27696.
 KW Hormone; Amidation.
 FT MOD_RES 6 6
 SQ SEQUENCE 6 AA; 621 MW; 72C9C6876DD81000 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 G 1
 Db 1 G 1

RESULT 43

FARP_MONEX
 ID FARP_MONEX STANDARD; PRT; 6 AA.
 AC P41966;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE FMRFamide-like neuropeptide GNFRF-amide.
 OS Moniezia expansa (Sheep tapeworm).
 OC Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
 OC Cyclophyllidae; Anoplocephalidae; Moniezia.
 OX NCBI_TaxID=28841;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=93312289; PubMed=8323531;
 RA Maule A.G., Shaw C., Halton D.W., Thim L.;
 RT "GNFRFamide: a novel FMRFamide-immunoreactive peptide isolated from

RT the sheep tapeworm, *Moniezia expansa*.
 RL Biochem. Biophys. Res. Commun. 193:1054-1060 (1993).
 RC -1- SIMILARITY: BELONGS TO THE FARP (FMRP-RELATED PEPTIDE) FAMILY.

CC PIR, A43129; A43129.
 DR Neuropeptide; Amidation.
 KW MOD RES 6 6
 FT SEQUENCE 6 AA; 787 MW; 69D409C9C4481000 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 DB 1 G 1

RESULT 44
 LOKI-LOCI
 ID LOKI-LOCI STANDARD; PRT; 6 AA.
 AC P4191;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Locustakinin I.

OS Locusta migratoria (Migratory locust).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;
 OC Acridoidea; Acrididae; Oedipodinae; Locusta.
 OC NCBI_TaxID=7004;
 RN [1]

SEQUENCE.
 RC TISSUE=Corpora cardiaca;
 RX MEDLINE=92262851; PubMed=1585017;
 RA Schoofs L., Holman G.M., Proost P., van Damme J., Hayes T.K.,
 de Loof A.;
 RT "Locustakinin, a novel myotropic peptide from Locusta migratoria,
 isolation, primary structure and synthesis";
 RL Regul. Pept. 37:49-57 (1992).

CC -1- FUNCTION: Myotropic peptide. May be important in the stimulation
 of ion transport and inhibition of diuretic activity in Malpighian
 tubules.

CC -1- SUBCELLULAR LOCATION: Secreted.

DR PIR, A61068; A61068.

GW Neuropeptide; Amidation.

MOD_RES 6 6
 FT SEQUENCE 6 AA; 654 MW; 686365A5B9CDB000 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 S 5
 DB 3 S 3

RESULT 45
 ALL2-CARMA
 ID ALL2-CARMA STANDARD; PRT; 7 AA.
 AC P81805;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Carcinustatin 2.

OS Carcinus maenas (Common shore crab) (Green crab).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
 OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
 OC NCBI_TaxID=6759;
 RN [1]

RP SEQUENCE.

RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
 RX MEDLINE=98121193; PubMed=9461295;
 RA Duvé H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
 Thorpe A.;
 RT "Isolation and identification of multiple neuropeptides of the
 RT allatostatin superfamily in the shore crab *Carcinus maenas*."
 RL Eur. J. Biochem. 250:727-734 (1997).
 CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 KW Neuropeptide; Amidation; Multigene family.
 FT MOD_RES 7 7
 FT SEQUENCE 7 AA; 770 MW; 672879CDB5DB70 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 DB 6 G 6

RESULT 46

ALL3-CARMA
 ID ALL3-CARMA STANDARD; PRT; 7 AA.

AC P81806;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Carcinustatin 3.
 OS Carcinus maenas (Common shore crab) (Green crab).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
 OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
 OC NCBI_TaxID=6759;
 RN [1]

SEQUENCE.

RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
 RX MEDLINE=98121193; PubMed=9461295;
 RA Duvé H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
 Thorpe A.;
 RT "Isolation and identification of multiple neuropeptides of the
 RT allatostatin superfamily in the shore crab *Carcinus maenas*."
 RL Eur. J. Biochem. 250:727-734 (1997).

CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.

CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.

KW Neuropeptide; Multigene family.

SQ SEQUENCE 7 AA; 796 MW; 672879CDB476B70 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 DB 6 G 6

RESULT 47

ALL4-CARMA
 ID ALL4-CARMA STANDARD; PRT; 7 AA.

AC P81807;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Carcinustatin 4.
 OS Carcinus maenas (Common shore crab) (Green crab).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
 OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
 OC NCBI_TaxID=6759;
 RN [1]

RP SEQUENCE.

RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
 RX MEDLINE=98121193; PubMed=9461295;
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
 RA Thorpe A.;
 RT "Isolation and identification of multiple neuropeptides of the
 RT allatostatin superfamily in the shore crab *Carcinus maenas*.";
 RL Eur. J. Biochem. 250:727-734(1997).
 CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 KW Neuropeptide; Multigene family.
 FT MOD RES 7
 SQ SEQUENCE 7 AA; 782 MW; 672879DCB476AC0 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 G 1
 Db 6 G 6

RESULT 48

ALL5_CARMA

ID ALL5_CARMA STANDARD; PRT; 7 AA.
 AC P81808;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Carcinustatin 5.
 OS *Carcinus maenas* (Common shore crab) (Green crab).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
 OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
 OX NCBI_TaxID=6759;
 RN [1]
 RP SEQUENCE
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
 RX MEDLINE=98121193; PubMed=9461295;
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
 RA Thorpe A.;
 RT "Isolation and identification of multiple neuropeptides of the
 RT allatostatin superfamily in the shore crab *Carcinus maenas*.";
 RL Eur. J. Biochem. 250:727-734(1997).
 CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 KW Neuropeptide; Amidation; Multigene family.
 FT MOD RES 7
 SQ SEQUENCE 7 AA; 781 MW; 672879DCB476420 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 G 1
 Db 6 G 6

RESULT 49

ALL7_CYDPO

ID ALL7_CYDPO STANDARD; PRT; 7 AA.
 AC P82158;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Cydiastatin 7.
 OS *Cydia pomonella* (Codling moth).
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Tortricidae; Tortricidae; Olethreutinae; Cydia.
 OX NCBI_TaxID=82600;
 RN [1]
 RP SEQUENCE.

RC TISSUE=Larva;
 RX MEDLINE=98054539; PubMed=9392829;
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D.,
 RA Davey M., East P.D., Thorpe A.;
 RT "Lepidopteran peptides of the allatostatin superfamily.";
 RL Peptides 18:1301-1309(1997).
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 KW Neuropeptide; Amidation.
 FT MOD RES 7
 SQ SEQUENCE 7 AA; 873 MW; 672879CABB569350 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 G 1
 Db 6 G 6

RESULT 50

BRHP_CONIM

ID BRHP_CONIM STANDARD; PRT; 7 AA.
 AC P58803;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Broncheptapeptide Im.
 OS *Conus imperialis* (Imperial cone).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
 OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
 OC Neogastropoda; Conoidea; Conidae; Conus.
 OX NCBI_TaxID=35631;
 RN [1]
 RP SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.
 RC TISSUE=Venom;
 RX MEDLINE=97184108; PubMed=9030520;
 RA Craig A.G., Jimenez E.C., Dykert J., Nielsen D.B., Gulyas J.,
 RA Abogadie F.C., Porter J., Rivier J.E., Cruz L.J., Olivera B.M.,
 RA McIntosh J.M.;
 RT "A novel post-translational modification involving bromination of
 RT tryptophan. Identification of the residue, L-6-bromotryptophan, in
 RT peptides from *Conus imperialis* and *Conus radiatus* venom.";
 RL J. Biol. Chem. 272:4689-4698(1997).
 CC -!- FUNCTION: Does not elicit gross behavioral symptoms when injected
 CC centrally or peripherally in mice.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
 CC -!- MASS SPECTROMETRY: MW=853.19; METHOD=LSIMS.
 DR PIR; A58512; A58512.
 KW Bromination; Amidation; Pyrrolidone carboxylic acid.
 FT DISULFID 2 7
 FT MOD RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD RES 6 6 BROMINATION.
 FT MOD RES 7 7 AMIDATION.
 SQ SEQUENCE 7 AA; 795 MW; 6EA37DC6D87EA6B0 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 G 1
 Db 3 G 3

RESULT 51

CHOX_ALCSP

ID CHOX_ALCSP STANDARD; PRT; 7 AA.
 AC F16101;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 01-APR-1990 (Rel. 14, Last annotation update)

DE Choline oxidase (EC 1.1.3.17) (Fragment).
 DS Alcaligenes sp.
 CC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 CC Alcaligenaceae; Alcaligenes.
 CC NCBI_TaxID=512;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=81006769; PubMed=6997283;
 RA Ohta-Fukuyama M., Miyake Y., Emi S., Yamano T.;
 RT "Identification and properties of the prosthetic group of choline
 oxidase from *Alcaligenes* sp.";
 RL J. Biochem. 88:197-203(1980).
 CC -1- CATALYTIC ACTIVITY: Choline + O(2) = betaine aldehyde + H(2)O(2).
 CC PIR; A15398; A15398.
 CC Oxidoreductase.
 CC NON TER 7
 ST SEQUENCE 7 AA; 839 MW; 7415B1E457644AC0 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 S 5
 |
 DB 6 S 6

RESULT 52
 CIA_ENTFA
 ID - CIA_ENTFA STANDARD; PRT; 7 AA.
 AC P11932;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-FEB-1991 (Rel. 17, Last annotation update)
 DE Sex pheromone CAM373 (Clumping-inducing agent) (CIA).
 DE Enterococcus faecalis (Streptococcus faecalis).
 CC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
 CC NCBI_TaxID=1351;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=87005252; PubMed=3093276;
 RA Mori M., Tanaka H., Sakagami Y., Isogai A., Fujino M., Kitada C.,
 RA White B.A., An F.Y., Clewell D.B., Suzuki A.;
 RT "Isolation and structure of the Streptococcus faecalis sex pheromone,
 CAM373";
 RL FEBS Lett. 206:69-72(1986).
 CC -1- FUNCTION: CAM373 INDUCES MATING RESPONSE OF DONOR CELLS
 CC HARBORING PAM373.
 CC -1- MISCELLANEOUS: THE N-TERMINUS IS POSSIBLY RESPONSIBLE FOR
 CC SPECIFICITY OF PHEROMONES TO PLASMIDS.
 CC -1- SIMILARITY: C-TERMINAL TO THE PHEROMONES CPD1 AND CAD1.
 CC PIR; A25269; A25269.
 CC Pheromone.
 ST SEQUENCE 7 AA; 734 MW; 75BDD72059C05DB0 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 S 5
 |
 DB 7 S 7

RESULT 53
 FAR1_HELTI
 ID - FAR1_HELTI STANDARD; PRT; 7 AA.
 AC P41871;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DE FMRFamide-like neuropeptide GDPPLRF-amide.
 CC Helisoma trivolvis (Snail).

OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
 OC Lymnaeidae; Planorbidae; Helisoma.
 CC NCBI_TaxID=27815;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=94286417; PubMed=7912428;
 RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;
 RT "FMRFamide-related peptides from the kidney of the snail, *Helisoma*
 trivolvis";
 RL Peptides 15:31-36(1994).
 CC -1- FUNCTION: APPEARS TO BE INVOLVED IN OSMOREGULATION BY AFFECTING
 CC THE KIDNEY, MANTLE AND SKIN.
 CC -1- TISSUE SPECIFICITY: KIDNEY, SKIN, MANTLE AND THE HEMOLYMPH.
 CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
 CC FAMILY.
 CC Neuropeptide; Amidation.
 CC MOD RES 7
 ST SEQUENCE 7 AA; 851 MW; 69D40729D76AA810 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 |
 DB 1 G 1

RESULT 54
 FAR3_HAECO
 ID - FAR3_HAECO STANDARD; PRT; 7 AA.
 AC P81258;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE FMRFamide-like neuropeptide PF3 (KSAYMRF-amide).
 CC Haemochus contortus (Barber pole worm).
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
 CC Trichostrongyloidea; Haemonchidae; Haemonchinae; Haemonchus.
 CC NCBI_TaxID=6289;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=99318264; PubMed=10391380;
 RA Marks N.J., Sangster N.C., Maule A.G., Halton D.W., Thompson D.P.,
 RA Geary T.G., Shaw C.;
 RT "Structural characterisation and pharmacology of KHEYLRFamide (AF2)
 and KSAYMRFamide (PF3/AF8) from *Haemochus contortus*";
 RL Mol. Biochem. Parasitol. 100:185-194(1999).
 CC -1- FUNCTION: ACTIVE ON NEUROMUSCULATURE.
 CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
 CC FAMILY.
 CC Neuropeptide; Amidation.
 CC MOD RES 7
 ST SEQUENCE 7 AA; 902 MW; 69D4068B5DC5B350 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 S 5
 |
 DB 2 S 2

RESULT 55
 FAR3_PANRE
 ID - FAR3_PANRE STANDARD; PRT; 7 AA.
 AC P41874;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DE FMRFamide-like neuropeptide GDPPLRF-amide.
 CC Helisoma trivolvis (Snail).

DE FMRamide-like neuropeptide PF3 (KSAYMRF-amide).
 OS Panagrellus redivivus.
 OC Eukaryota; Metazoa; Chromadorea; Rhabditida;
 OC Panagrolaimoidea; Panagrolaimidae; Panagrellus.
 OX NCBI_TaxID=6233;
 RN [1]
 RP SEQUENCE, AND SYNTHESIS.
 RX MEDLINE=94235053; PubMed=8179635;
 RA Maule A.G., Shaw C., Bowman J.W., Halton D.W., Thompson D.P.,
 RA Geary T.G., Thim L.;
 RT "KSAYMRFamide: a novel FMRamide-related heptapeptide from the free-
 RT living nematode, Panagrellus redivivus, which is myoactive in the
 RT parasitic nematode, Ascaris suum.";
 RL Biochem. Biophys. Res. Commun. 200:973-980(1994).
 CC -!- FUNCTION: MYOACTIVE; INDUCES A RAPID CONCENTRATION-DEPENDENT
 CC MUSCLE TENSION INCREASE.
 CC -!- SIMILARITY: BELONGS TO THE FARP (FMRAMIDE RELATED PEPTIDE)
 CC FAMILY.
 DR PIR; PC2132.
 KW Neuropeptide; Amidation.
 FT MOD RES 7
 SQ SEQUENCE 7 AA; 902 MW; 69D4068B5DC5B350 CRC64;
 AMIDATION.
 Query Match 20.0%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 S 5
 DB 2 S 2
 RESULT 56
 GFRP MOUSE
 ID _GFRP_MOUSE STANDARD; PRT; 7 AA.
 AC P99025;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE GTP cyclohydrolase I feedback regulatory protein (P35) (Fragment).
 GN GCHFR OR GFRP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Liver;
 RA Sanchez J.-C., Rouge V., Frutiger S., Hughes G., Yan J.X.,
 RA Hoogland C., Appel R.D., Binz P.-A., Hochstrasser D.F.,
 RA Cowthorne M.;
 RA Submitted (AUG-1998) to the SWISS-PROT data bank.
 CC -!- FUNCTION: MEDIATES TETRAHYDROBIPTERIN INHIBITION OF GTP
 CC CYCLOHYDROLASE I. THIS INHIBITION IS REVERSED BY L-PHENYLALANINE
 CC (BY SIMILARITY).
 CC -!- SUBUNIT: Homodimer (By similarity).
 DR SWISS-2DPAGE; P99025; MOUSE.
 FT INIT MET 0
 FT NON TER 7
 SQ SEQUENCE 7 AA; 806 MW; 71B5B057273B4700 CRC64;
 Query Match 20.0%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 S 5
 DB 6 S 6
 RESULT 57
 HY7_PIG
 ID _HY7_PIG STANDARD; PRT; 7 AA.

AC P01153;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 21-JUL-1986 (Rel. 01, Last annotation update)
 DE Hypothalamic heptapeptide.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE, AND SYNTHESIS.
 RX MEDLINE=81213980; PubMed=6263778;
 RA Chang R.C.C., Huang W.-Y., Arimura A., Redding T.W., Coy D.H.,
 RA Saffran M., Kong A., Hamilton J.W., Cohn D.V., Schally A.V.;
 RT "Isolation, structure and synthesis of a heptapeptide with in vitro
 RT ACTH-releasing activity from porcine hypothalamus.";
 RL Horm. Metab. Res. 13:228-232(1981).
 DR PIR; A01417; NYPG7.
 SQ SEQUENCE 7 AA; 957 MW; 632B45B1FB5059A0 CRC64;
 Query Match 20.0%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 S 5
 DB 5 S 5
 RESULT 58
 IGAO_DACDE
 ID IGAO_DACDE STANDARD; PRT; 7 AA.
 AC P06294;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)
 DE Galactose oxidase inhibitor.
 OS Dactylium dendroides (Cladobotryum dendroides).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypomyces.
 OX NCBI_TaxID=5132;
 RN [1]
 RP SEQUENCE.
 RA Avigad G., Markus Z.;
 RT "Identification of a peptide inhibitor of galactose oxidase from
 RT Dactylium dendroides.";
 RL Fed. Proc. 31:447-447(1972).
 CC -!- FUNCTION: BINDS ONE COPPER ION PER MOLECULE BUT DOES NOT BIND THE
 CC GALACTOSE OXIDASE APOENZYME. IT MAY INACTIVATE THE ENZYME BY
 CC BINDING TO ITS PROSTHETIC COPPER GROUP.
 DR PIR; A01341; XEYDGD.
 KW Copper; Metalloenzyme inhibitor.
 SQ SEQUENCE 7 AA; 706 MW; 75B801A456D87DB0 CRC64;
 Query Match 20.0%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 G 1
 DB 2 G 2
 RESULT 59
 MNPI_LEPDE
 ID _MNPI_LEPDE STANDARD; PRT; 7 AA.
 AC P42984;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Myotropic neuropeptide 1 (led-MNP-I).
 OS Leptinotarsa decemlineata (Colorado potato beetle).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
 OC Phycophaga; Chrysomeloidea; Chrysomelidae; Chrysomelinae;
 OC Chrysomelini; Leptinotarsa.
 OX NCBI_TaxID=7539;
 [1]
 RN
 SEQUENCE, AND SYNTHESIS.
 RC TISSUE=Head;
 RX MEDLINE=95380343; PubMed=7651886;
 RA Spittaels L., Vankeerberghen A., Schoofs L., Torrekens S.,
 RA Grauwels L., van Leuven F., de Loof A.;
 RT "Identification, characterization, and immunological localization of
 RT a novel myotropic neuropeptide in the Colorado potato beetle,
 RT Leptinotarsa decemlineata";
 RL Peptides 16:365-374(1995).
 CC -!- FUNCTION: MYOTROPIC PEPTIDE. STIMULATES THE CONTRACTIONS OF THE
 CC OVIDUCT.
 CC Neuropeptide; Amidation.
 KW MOD RES 7 7 AMIDATION.
 SQ SEQUENCE 7 AA; 705 MW; 6DD73768745B5DB0 CRC64;
 Query Match 20.0%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 G 1
 Db 4 G 4

RESULT 60
 PPH2 LYCES STANDARD; PRT; 7 AA.
 ID AC P83379;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Purple acid phosphatase isozyme LeSAP2 (EC 3.1.3.2) (Fragment).
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Asteridae; Lamiales; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4081;
 [1]
 RN
 SEQUENCE, CATALYTIC ACTIVITY, SUBUNIT, SUBCELLULAR LOCATION, AND
 GP GLYCOSYLATION.
 RC STRAIN=cy. Money-maker; TISSUE=Seed;
 RX MEDLINE=22361242; PubMed=12473124;
 RA Bozzo G.G., Raghothama K.G., Plaxton W.C.;
 RT "Purification and characterization of two secreted purple acid
 RT phosphatase isozymes from phosphate-starved tomato (Lycopersicon
 RT esculentum) cell cultures";
 RL Eur. J. Biochem. 269:6278-6286(2002).
 CC -!- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an
 CC alcohol + phosphate.
 CC -!- SUBUNIT: Monomer.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- PTM: Glycosylated.
 CC -!- MISCELLANEOUS: In L. esculentum there are at least two isozymes of
 CC purple acid phosphatase.
 CC Hydrolase; Glycoprotein.
 KW FT NON TER 1 1
 FT NON TER 7 7
 SQ SEQUENCE 7 AA; 810 MW; 672AA862C9C729A0 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 G 1
 Db 5 G 5

RESULT 61
 UC24 MAIZE STANDARD; PRT; 7 AA.
 ID AC P80630;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Unknown protein from 2D-page of etiolated coleoptile (Spot 447)
 DE (Fragment).
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACAD clade; Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxID=4577;
 [1]
 RN
 SEQUENCE.
 RC TISSUE=Coleoptile;
 RA Tounzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
 RA Pernellet J.-C., Zivy M., de Vienne D.;
 RT "The maize two dimensional gel protein database: towards an integrated
 RT genome analysis program";
 RL Theor. Appl. Genet. 93:997-1005(1996).
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 6.0, ITS MW IS: 30.0 kDa.
 DR Maize-2DPAGE; P80630; COLEOPTILE.
 DR MaizeDB; 123956; - 1
 FT NON TER 7 7
 FT NON TER 1 1
 SQ SEQUENCE 7 AA; 665 MW; 6DC1B5B33DC1B5D0 CRC64;
 Query Match 20.0%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 5 S 5
 Db 1 S 1

RESULT 62
 UF04 MOUSE STANDARD; PRT; 7 AA.
 ID AC P38642;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE Unknown protein from 2D-page of fibroblasts (P46) (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 [1]
 RN
 SEQUENCE.
 RC TISSUE=Fibroblast;
 RX MEDLINE=95009907; PubMed=7523108;
 RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
 RT "Separation and sequencing of familial and novel murine proteins
 RT using preparative two-dimensional gel electrophoresis";
 RL Electrophoresis 15:735-745(1994).
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 5.0, ITS MW IS: 46 kDa.
 FT NON TER 7 7
 FT NON TER 1 1
 SQ SEQUENCE 7 AA; 766 MW; 68640AB77632700 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 G 1
 Db 7 G 7

RESULT 63

UH11_RAT
ID_UH11_RAT STANDARD; PRT; 7 AA.
AC P56576;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Unknown protein from 2D-page of heart tissue (Spot P11) (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
ON NCBI_TaxID=10116;
RN [1]
RP SEQUENCE.
RC STRAIN=Wistar; TISSUE=Heart;
RA Li X.-P., Pleissner K.-P., Scheler C., Regitz-Zagrosek V., Salikov J.,
RA Jungblut P.R.;
RL Submitted (SEP-1998) to the SWISS-PROT data bank.
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 8.5, ITS MW IS: 42 kDa.
FT UNSURE 2 2 OR A.
FT NON TER 7
SQ SEQUENCE 7 AA; 775 MW; 6866DB040DC5A6B0 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
DB 7 G 7

RESULT 64

UN06_PINPS
ID_UN06_PINPS STANDARD; PRT; 7 AA.
AC P81675;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Unknown protein from 2D-page of needles (N141) (Fragment).
OS Pinus pinaster (Maritime pine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
ON NCBI_TaxID=71647;
RN [1]
RP SEQUENCE.
RC TISSUE=Needle;
RX MEDLINE=99274088; PubMed=10344291;
RA Costa P., Pionneau C., Bauw G., Dubos C., Bahrman N., Kremer A.,
RA Frigerio J.-M., Plomion C.;
RT "Separation and characterization of needle and xylem maritime pine
RT proteins".
RL Electrophoresis 20:1098-1108(1999).
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 6.6, ITS MW IS: 25 kDa.
FT NON TER 1 1
FT NON TER 7
SQ SEQUENCE 7 AA; 823 MW; 69D76724486B5740 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
DB 3 G 3

RESULT 65

WWA1_ACHF
ID_WWA1_ACHF STANDARD; PRT; 7 AA.
AC P35919;

DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Wwamide-1.
OS Achatina fulica (Giant African snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC Sigmurethra; Achatinoidea; Achatinidae; Achatina.
ON NCBI_TaxID=6530;
RN [1]
RP SEQUENCE.
RC TISSUE=Ganglion;
RX MEDLINE=93265912; PubMed=8495720;
RA Minakata H., Ikeda T., Muneoka Y., Kobayashi M., Nomoto K.;
RT "Wwamide-1, -2 and -3: novel neuromodulatory peptides isolated from
RT ganglia of the African giant snail, Achatina fulica".
RL FEBS Lett. 323:104-108(1993).
CC -!- FUNCTION: EXHIBITS MODULATORY EFFECTS ON THE PERIPHERAL NERVOUS
CC SYSTEM. INHIBITS ACTIVITY ON A CENTRAL NEURON.
DR PIR; S33245; S33245.
KW Neuropeptide; Amidation.
FT MOD.RES 7
FT SEQUENCE 7 AA; 993 MW; 7362D5B69B041310 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 S 5
DB 5 S 5

RESULT 66

WWA2_ACHF
ID_WWA2_ACHF STANDARD; PRT; 7 AA.
AC P35920;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Wwamide-2.
OS Achatina fulica (Giant African snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC Sigmurethra; Achatinoidea; Achatinidae; Achatina.
ON NCBI_TaxID=6530;
RN [1]
RP SEQUENCE.
RC TISSUE=Ganglion;
RX MEDLINE=93265912; PubMed=8495720;
RA Minakata H., Ikeda T., Muneoka Y., Kobayashi M., Nomoto K.;
RT "Wwamide-1, -2 and -3: novel neuromodulatory peptides isolated from
RT ganglia of the African giant snail, Achatina fulica".
RL FEBS Lett. 323:104-108(1993).
DR PIR; S33246; S33246.
KW Neuropeptide; Amidation.
FT MOD.RES 7
FT SEQUENCE 7 AA; 964 MW; 7362D5B686D32310 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 S 5
DB 5 S 5

RESULT 67

WWA3_ACHF
ID_WWA3_ACHF STANDARD; PRT; 7 AA.
AC P35921;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)


```

DE Wamide-3.
OS Achatina fulica (Giant African snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC Sigmurethra; Achatinoidea; Achatinidae; Achatina.
OX NCBI_TaxID=6530;
RN [1]
RP SEQUENCE.
RC TISSUE=Ganglion;
RC MEDLINE=93265912; PubMed=8495720;
RA Minakata H., Ikeda T., Muneoka Y., Kobayashi M., Nomoto K.;
RT "Wamide-1, -2 and -3: novel neuromodulatory peptides isolated from
RT ganglia of the African giant snail, Achatina fulica.";
RL FEBS Lett. 323:104-108(1993).
RL PIR; S33244; S33244.
DR Neuropeptide; Amidation.
KW MOD_RES 7 7
FT SEQUENCE 7 AA; 965 MW; 7362D5B69B132310 CRC64;
SQ
Query Match 20.0%; Score 1; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 S 5
DB 5 S 5

RESULT 68
ACI_THUAL
ID -ACI_THUAL STANDARD; PRT; 8 AA.
AC P18691;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-NOV-1990 (Rel. 16, Last annotation update)
DE Angiotensin-converting enzyme inhibitor.
DE Thunus albacares (Yellowfin tuna) (Neothunnus macropterus).
JC Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Ruteleostomi;
JC Actinopterygii; Neopterygii; Teleostei; Ruteleostei; Neoteleostei;
JC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
JC Scombridae; Thunnus.
CX NCBI_TaxID=8236;
CX [1]
XP SEQUENCE.
XC TISSUE=Muscle;
XC MEDLINE=88326322; PubMed=3415688;
XA Kohana Y., Matsumoto S., Oka H., Teramoto T., Okabe M., Mimura T.;
RT "Isolation of angiotensin-converting enzyme inhibitor from tuna
RT muscle.";
XL Biochem. Biophys. Res. Commun. 155:332-337(1988).
DR PIR; A31570; A31570.
SQ SEQUENCE 8 AA; 953 MW; 6AA863733051F1B7 CRC64;
Query Match 20.0%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
ZY 1 G 1
ZB 7 G 7

RESULT 69
AKHG GRBYI
ID -AKHG GRBYI STANDARD; PRT; 8 AA.
AC P14086;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Adipokinetic hormone G (AKH-G) (RO II).
DE Gryllus bimaculatus (Two-spotted cricket), and
DE Romalea microptera (Lubber grasshopper).
JC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
JC Neoptera; Orthopteroidea; Orthoptera; Ensifera; Gryllinae;

```

```

OC Gryllus.
OX NCBI_TaxID=6999, 7007;
RN [1]
RP SEQUENCE.
RC SPECIES=G.bimaculatus; TISSUE=Corpora cardiaca;
RX MEDLINE=88106553; PubMed=3426616;
RA Gaede G., Rinehart K.L. Jr.;
RT "Primary sequence analysis by fast atom bombardment mass spectrometry
RT of a peptide with adipokinetic activity from the corpora cardiaca of
RL the cricket Gryllus bimaculatus.";
RL Biochem. Biophys. Res. Commun. 149:908-914(1987).
RN [2]
RP SEQUENCE.
RC SPECIES=R.microptera; TISSUE=Corpora cardiaca;
RX MEDLINE=89145002; PubMed=3226948;
RA Gaede G., Hilblisch C., Beyreuther K., Rinehart K.L. Jr.;
RT "Sequence analyses of two neuropeptides of the AKH/RPCH-family from
RT the lubber grasshopper, Romalea microptera.";
RL Peptides 9:681-688(1988).
CC -1- FUNCTION: THIS HORMONE, RELEASED FROM CELLS IN THE CORPORA
CC CARDIACA AFTER THE BEGINNING OF FLIGHT, CAUSES RELEASE OF
CC DIGLYCERIDES FROM THE FAT BODY AND THEN STIMULATES THE FLIGHT
CC MUSCLES TO USE THESE DIGLYCERIDES AS AN ENERGY SOURCE.
CC -1- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR PIR; A28004; A28004.
DR InterPro; IPR002047; AKH.
DR PROSITE; PS00256; AKH; 1.
KW Neuropeptide; Amidation; Flight; Pyrrolidone carboxylic acid.
FT MOD_RES 1 1
FT MOD_RES 8 8
FT MOD_RES 8 8
SQ SEQUENCE 8 AA; 938 MW; 867861B5B9C452D6 CRC64;
Query Match 20.0%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 S 5
DB 5 S 5

RESULT 70
AKH_LIBAU
ID -AKH_LIBAU STANDARD; PRT; 8 AA.
AC P25418;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Adipokinetic hormone (AKH).
OS Libellula auripennis (Skimmer dragonfly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Palaeoptera; Odonata; Anisoptera; Libellulidae; Libellula.
OX NCBI_TaxID=6965;
RN [1]
RP SEQUENCE. AND SYNTHESIS.
RC TISSUE=Corpora cardiaca;
RX MEDLINE=90359055; PubMed=2390213;
RA Gaede G.;
RT "The putative ancestral peptide of the adipokinetic/red-pigment-
RT concentrating hormone family isolated and sequenced from a
RT dragonfly.";
RL Biol. Chem. Hoppe-Seyler 371:475-483(1990).
CC -1- FUNCTION: THIS HORMONE, RELEASED FROM CELLS IN THE CORPORA
CC CARDIACA AFTER THE BEGINNING OF FLIGHT, CAUSES RELEASE OF
CC DIGLYCERIDES FROM THE FAT BODY AND THEN STIMULATES THE FLIGHT
CC MUSCLES TO USE THESE DIGLYCERIDES AS AN ENERGY SOURCE.
CC -1- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR PIR; S10596; S10596.
DR InterPro; IPR002047; AKH.
DR PROSITE; PS00256; AKH; 1.
KW Neuropeptide; Amidation; Flight; Pyrrolidone carboxylic acid.
FT MOD_RES 1 1
FT MOD_RES 8 8
FT MOD_RES 8 8

```

SQ SEQUENCE 8 AA; 978 MW; 865A771A9C452D6 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 S 5
|
7 S 7

Db

RESULT 71
AKH MELML
ID AKH MELML STANDARD; PRT; 8 AA.

AC P25423;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Adipokinetic hormone (AKH)
OS Melolontha melolontha (Cockchafer), and
OS Geotrupes stercorosus (Dor beetle), and
OS Pachnoda marginata (flower beetle).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Scarabaeiformia;
OC Scarabaeidae; Melolonthinae; Melolontha.
OX NCBI_TaxID=7061, 7087, 7058;
RN [1]
RP SEQUENCE.
RC SPECIES=M. melolontha, and G. stercorosus; TISSUE=Corpora cardiaca;
RX MEDLINE=91248100; PubMed=2039445;
RA Gaede G.;
RT "A unique charged tyrosine-containing member of the adipokinetic
RT hormone/red-pigment-concentrating hormone peptide family isolated and
RT sequenced from two beetle species.";
RL Biochem. J. 275:671-677(1991).
RN [2]

RP SEQUENCE.
RC SPECIES=P. marginata; TISSUE=Corpora cardiaca;
RX MEDLINE=92265187; PubMed=1586453;
RA Gaede G., Lopata A., Kellner R., Rinehart K.L. Jr.;
RT "Primary structures of neuropeptides isolated from the corpora
RT cardiaca of various cetonid beetle species determined by
RT pulsed-liquid phase sequencing and tandem fast atom bombardment mass
RT spectrometry.";
RL Biol. Chem. Hoppe-Seyler 373:133-142(1992).
CC -!- FUNCTION: THIS HORMONE, RELEASED FROM CELLS IN THE CORPORA
CC CARDIACA AFTER THE BEGINNING OF FLIGHT, CAUSES RELEASE OF
CC DIGLYCERIDES FROM THE FAT BODY AND THEN STIMULATES THE FLIGHT
CC MUSCLES TO USE THESE DIGLYCERIDES AS AN ENERGY SOURCE.
CC -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR PIR; A58641; A58641.
DR PIR; S15422; S15422.
DR InterPro: IPR002047; AKH.
DR PROSITE: PS00256; AKH; 1.
KW Neuropeptide; Amidation; Flight; Pyrrolidone carboxylic acid.
FT MOD RES 1 1
FT MOD RES 8 8
FT MOD RES 8 8
SQ SEQUENCE 8 AA; 1022 MW; 867AB775AB544736 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 S 5
|
5 S 5

Db

RESULT 72
AKH TABAT
ID AKH TABAT STANDARD; PRT; 8 AA.
AC P14595;
DT 01-JAN-1990 (Rel. 13, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Adipokinetic hormone (AKH) (Dipteran corpora cardiaca factor 1)
DE (DCC 1).
OS Tabanus atratus (Horse fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Tabanomorpha; Tabanidae;
OC Tabanus.
OX NCBI_TaxID=7207;
RN [1]
RP SEQUENCE.
RC TISSUE=Corpora cardiaca;
RX MEDLINE=90046758; PubMed=2813385;
RA Jaffe H., Raina A.K., Riley C.T., Fraser B.A., Nachman R.J.,
RA Vogel V.W., Zhang Y.-S., Hayes D.K.;
RT "Primary structure of two neurotrophic hormones with adipokinetic and
RT hypotrehalosemic activity isolated from the corpora cardiaca of horse
RT flies (Diptera).";
RL Proc. Natl. Acad. Sci. U.S.A. 86:8161-8164 (1989).
CC -!- FUNCTION: THIS HORMONE, RELEASED FROM CELLS IN THE CORPORA
CC CARDIACA AFTER THE BEGINNING OF FLIGHT, CAUSES RELEASE OF
CC DIGLYCERIDES FROM THE FAT BODY AND THEN STIMULATES THE FLIGHT
CC MUSCLES TO USE THESE DIGLYCERIDES AS AN ENERGY SOURCE.
CC -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR PIR; A33995; A33995.
DR InterPro: IPR002047; AKH.
DR PROSITE: PS00256; AKH; 1.
KW Neuropeptide; Amidation; Flight; Pyrrolidone carboxylic acid.
FT MOD RES 1 1
FT MOD RES 8 8
FT MOD RES 8 8
SQ SEQUENCE 8 AA; 949 MW; 86786771A9D1A736 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
|
7 G 7

Db

RESULT 73
AL12-CARMA
ID AL12-CARMA STANDARD; PRT; 8 AA.
AC P81815;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 12.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunioidea; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Multigene family.
SQ SEQUENCE 8 AA; 913 MW; 672879CDB569AB7 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1

```

Db          7 G 7
|
5 S 5
1 S 1

RESULT 74
AL15_CARMA STANDARD; PRT; 8 AA.
ID AL15_CARMA
AC P81818;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 15.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation; Multigene family.
FT MOD RES 8 8
SQ SEQUENCE 8 AA; 811 MW; 922879D5AB47687D CRC64;

Query Match 20.0%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred.No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y 1 G 1
|
2 G 2

RESULT 75
AL17_CARMA STANDARD; PRT; 8 AA.
ID AL17_CARMA
AC P81820;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 17.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation; Multigene family.
FT MOD RES 8 8
SQ SEQUENCE 8 AA; 858 MW; C82879D5AB46D865 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred.No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y 1 G 1
|
2 G 2

RESULT 76
AL18_CARMA STANDARD; PRT; 8 AA.
ID AL18_CARMA
AC P81821;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 18.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation; Multigene family.
FT MOD RES 8 8
SQ SEQUENCE 8 AA; 919 MW; C82879D5AB569AB5 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred.No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y 1 G 1
|
7 G 7

RESULT 77
AL11_CYPDO STANDARD; PRT; 8 AA.
ID AL11_CYPDO
AC P82152;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cydiastatin 1.
OS Cydia pomonella (Codling moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Tortricidae; Tortricidae; Olethreutinae; Cydia.
OX NCBI_TaxID=82600;
RN [1]
RP SEQUENCE.
RC TISSUE=Larva;
RX MEDLINE=98054539; PubMed=9392829;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D.,
RA Davey M., East P.D., Thorpe A.;
RT "Lepidopteran peptides of the allatostatin superfamily.";
RL Peptides 18:1301-1309(1997).
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation.
FT MOD RES 8 8
SQ SEQUENCE 8 AA; 934 MW; C82879C45B51F775 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred.No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y 1 G 1
|
7 G 7

```

```

Db      7 G 7

RESULT 78
ALL3_CVDPO
ID _ALL3_CVDPO STANDARD; PRT; 8 AA.
AC P82154; 2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cydiastatin 3.
OS Cydia pomonella (Coddling moth).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Tortricidae; Tortricidae; Olethreutinae; Cydia.
OX NCBI_TaxID=82600;
RN [1]
RP SEQUENCE.
RC TISSUE=Larva;
RX MEDLINE=98054539; PubMed=9392829;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D.,
RA Davey M., East P.D., Thorpe A.;
RT "Lepidopteran peptides of the allatostatin superfamily.";
RL Peptides 18:1301-1309(1997).
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation.
FT MOD RES 8
SQ SEQUENCE 8 AA; 926 MW; C82879D5AB477415 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 G 1
Db      7 G 7

RESULT 79
ALL4_CALVO
ID _ALL4_CALVO STANDARD; PRT; 8 AA.
AC P41840;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Callatostatin 4 (Leu-callatostatin 4).
OS Calliphora vomitoria (Blue blowfly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
OX NCBI_TaxID=27454;
RN [1]
RP SEQUENCE.
RC TISSUE=Thoracic ganglion;
RX MEDLINE=93211980; PubMed=8460157;
RA Duve H., Johnsen A.H., Scott A.G., Yu C.G., Yagi K.J., Tobé S.S.,
RA Thorpe A.;
RT "Callatostatins: neuropeptides from the blowfly Calliphora vomitoria
with sequence homology to cockroach allatostatins.";
RL proc. Natl. Acad. Sci. U.S.A. 90:2456-2460(1993).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=94291167; PubMed=8020069;
RA Duve H., Thorpe A.;
RT "Distribution and functional significance of Leu-callatostatins in
the blowfly Calliphora vomitoria.";
RL Cell Tissue Res. 276:367-379(1994).
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR AND PLAY
A ROLE IN THE INTEGRATION OF INFORMATION WITHIN THE BRAIN. MAY BE
INVOLVED IN THE CONTROL OF VISCERAL MUSCLES DUE TO ITS ABILITY TO
BEHAVE AS POTENT INHIBITORS OF PERISTALTIC MOVEMENTS. MAY ALSO
FULFILL A NEUROHORMONAL ROLE ON MUSCLES OF THE GUT AND HEART.

-!- TISSUE SPECIFICITY: BRAIN, SUBESOPHAGEAL GANGLION, RETROCEBRAL
COMPLEX, THORACICO-ABDOMINAL GANGLION, PERIPHERAL NEUROSECRETORY
SYSTEM AND INTESTINE.
-!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation.
FT MOD RES 8
FT UNSURE 1 8 OR N.
SQ SEQUENCE 8 AA; 954 MW; D32879D5AB47740A CRC64;

Query Match 20.0%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      5 S 5
Db      5 S 5

RESULT 80
ALL4_CVDPO
ID _ALL4_CVDPO STANDARD; PRT; 8 AA.
AC P82155; 2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cydiastatin 4.
OS Cydia pomonella (Coddling moth).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Tortricidae; Tortricidae; Olethreutinae; Cydia.
OX NCBI_TaxID=82600;
RN [1]
RP SEQUENCE.
RC TISSUE=Larva;
RX MEDLINE=98054539; PubMed=9392829;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D.,
RA Davey M., East P.D., Thorpe A.;
RT "Lepidopteran peptides of the allatostatin superfamily.";
RL Peptides 18:1301-1309(1997).
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation.
FT MOD RES 8
SQ SEQUENCE 8 AA; 910 MW; 922879D5AB47740D CRC64;

Query Match 20.0%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      5 S 5
Db      5 S 5

RESULT 81
ALL5_CALVO
ID _ALL5_CALVO STANDARD; PRT; 8 AA.
AC P41841;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Callatostatin 5 (Met-callatostatin 1) (Hyp3)Met-callatostatin).
OS Calliphora vomitoria (Blue blowfly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
OX NCBI_TaxID=27454;
RN [1]
RP SEQUENCE.
RX MEDLINE=93211980; PubMed=8460157;
RA Duve H., Johnsen A.H., Scott A.G., Yu C.G., Yagi K.J., Tobé S.S.,
RA Thorpe A.;
RT "Callatostatins: neuropeptides from the blowfly Calliphora vomitoria

```

RT with sequence homology to cockroach allatostatins.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:2456-2460(1993).
 RN [2]
 RN CHARACTERIZATION, AND HYDROXYLATION.
 RC TISSUE=Head;
 RX MEDLINE=94342269; PubMed=8063725;
 DX Duve H., Johnsen A.H., Scott A.G., East P., Thorpe A.;
 RT "[Hyp3]Met-callatostatin. Identification and biological properties of
 RT a novel neuropeptide from the blowfly Calliphora vomitoria.";
 RL J. Biol. Chem. 269:21059-21066(1994).
 CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR AND PLAY
 CC A ROLE IN THE INTEGRATION OF INFORMATION WITHIN THE BRAIN. MAY BE
 CC INVOLVED IN THE CONTROL OF VISCERAL MUSCLES DUE TO ITS ABILITY TO
 CC BEHAVE AS POTENT INHIBITORS OF PERISTALTIC MOVEMENTS. MAY ALSO
 CC FULFILL A NEUROHORMONAL ROLE ON MUSCLES OF THE GUT AND HEART.
 CC -1- TISSUE SPECIFICITY: NEURONS WITHIN BRAIN AND ABDOMINAL GANGLION.
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 DR PIR: E47393; E47393.
 KW Neuropeptide; Amidation; Hydroxylation.
 KW MOD_RES 3 8 HYDROXYLATION (20%).
 FT MOD_RES 8 8 AMIDATION.
 SQ SEQUENCE 8 AA; 883 MW; 7D9879CABB477768 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 |
 DB 1 G 1

RESULT 82
 ALL5_CYPDPO STANDARD; PRT; 8 AA.
 ID AC P82156;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Cydiastatin 5.
 CS Cydia pomonella (Codling moth).
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 CC Tortricidae; Tortricidae; Olethreutinae; Cydia.
 CX NCBI_TaxID=82600;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Larva;
 RX MEDLINE=98054539; PubMed=9392829;
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D.,
 RA Davey M., East P.D., Thorpe A.;
 RT "Lepidopteran peptides of the allatostatin superfamily.";
 RL Peptides 18:1301-1309(1997)
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 CC Neuropeptide; Amidation.
 FT MOD_RES 8 8 AMIDATION.
 FT MOD_RES 8 8
 SQ SEQUENCE 8 AA; 898 MW; 922879CABB58640D CRC64;

Query Match 20.0%; Score 1; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 |
 DB 3 G 3

RESULT 83
 ALL6_CYPDPO STANDARD; PRT; 8 AA.
 ID AC P82157;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Cydiastatin 6.
 OS Cydia pomonella (Codling moth).
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 CC Tortricidae; Tortricidae; Olethreutinae; Cydia.
 CX NCBI_TaxID=82600;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Larva;
 RX MEDLINE=98054539; PubMed=9392829;
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D.,
 RA Davey M., East P.D., Thorpe A.;
 RT "Lepidopteran peptides of the allatostatin superfamily.";
 RL Peptides 18:1301-1309(1997).
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 CC Neuropeptide; Amidation.
 FT MOD_RES 8 8 AMIDATION.
 SQ SEQUENCE 8 AA; 936 MW; 0B2879C45B573767 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 |
 DB 7 G 7

RESULT 84
 ALL7_CARMA STANDARD; PRT; 8 AA.
 ID AC P81809;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Carcinustatin 7 [Contains: Carcinustatin 6; Carcinustatin 1].
 OS Carcinus maenas (Common shore crab) (Green crab).
 CC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 CC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
 CC Eubrachyura; Portunoidae; Portunidae; Carcinus.
 CX NCBI_TaxID=6759;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
 RX MEDLINE=98121193; PubMed=9461295;
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
 RA Thorpe A.;
 RT "Isolation and identification of multiple neuropeptides of the
 RT allatostatin superfamily in the shore crab Carcinus maenas.";
 RL Eur. J. Biochem. 250:727-734(1997).
 CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 CC Neuropeptide; Amidation; Multigene family.
 FT PEPTIDE 1 8 CARCINUSTATIN 7.
 FT PEPTIDE 2 8 CARCINUSTATIN 6.
 FT PEPTIDE 4 8 CARCINUSTATIN 1.
 FT MOD_RES 8 8 AMIDATION.
 SQ SEQUENCE 8 AA; 825 MW; 922879CDB4775BD CRC64;

Query Match 20.0%; Score 1; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
 |
 DB 7 G 7

RESULT 85
 ALL8_CARMA STANDARD; PRT; 8 AA.
 ID AC P81811;

DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Carcinustatin 8.
 OS Carcinus maenas (Common shore crab) (Green crab).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
 OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
 OX NCBI_TaxID=6759;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
 RX MEDLINE=98121193; PubMed=9461295;
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
 RA Thorpe A.;
 RT "Isolation and identification of multiple neuropeptides of the
 RT allatostatin superfamily in the shore crab Carcinus maenas.";
 RL Eur. J. Biochem. 250:727-734(1997).
 CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 KW Neuropeptide; Amidation; Multigene family.
 FT MOD_RES 8 8 AMIDATION.
 SQ SEQUENCE 8 AA; 795 MW; 922879CDBC847687D CRC64;

 Query Match 20.0%; Score 1; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 G 1
 DB 2 G 2

 RESULT 86
 CADI_ENTFA
 ID CADI_ENTFA STANDARD; PRT; 8 AA.
 AC P13268;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-FEB-1991 (Rel. 17, Last annotation update)
 DE Sex pheromone CADI.
 OS Enterococcus faecalis (Streptococcus faecalis).
 OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
 OX NCBI_TaxID=1351;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=85051889; PubMed=6437872;
 RA Mori M., Sagakami Y., Narita M., Isogai A., Fujino M., Kitada C.,
 RA Craig R.A., Clewell D.B., Suzuki A.;
 RT "Isolation and structure of the bacterial sex pheromone, cADI, that
 RT induces plasmid transfer in Streptococcus faecalis.";
 RL FEBS Lett. 178:97-100(1984).
 CC -!- FUNCTION: CADI IS INVOLVED IN THE CONJUGATIVE TRANSFER OF THE
 CC HEMOLYSIN PLASMID PAD1.
 KW Pheromone.
 SQ SEQUENCE 8 AA; 819 MW; 047DD732C735B9C7 CRC64;

 Query Match 20.0%; Score 1; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 G 1
 DB 8 G 8

 RESULT 87
 CCKX_MACEU
 ID CCKX_MACEU STANDARD; PRT; 8 AA.
 AC F30369;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Cholecystokinin (CCK).
 GN CCK.
 OS Macropus eugenii (Tamar wallaby), and
 OS Dasyurus viverrinus (Southeastern quoll).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
 OX NCBI_TaxID=9315, 9279;
 RN [1]
 RP SEQUENCE.
 RC SPECIES=M.eugenii, and D.viverrinus;
 RC TISSUE=Brain;
 RX MEDLINE=89234141; PubMed=3375140;
 RA Fan Z.W., Eng J., Shaw G., Yalow R.S.;
 RA "Cholecystokinin octapeptide purified from brains of Australian
 RT marsupials.";
 RL Peptides 9:429-431(1988).
 CC -!- FUNCTION: THIS PEPTIDE HORMONE INDUCES GALL BLADDER CONTRACTION
 CC AND THE RELEASE OF PANCREATIC ENZYMES IN THE GUT. ITS FUNCTION
 CC IN THE BRAIN IS NOT CLEAR.
 CC -!- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.
 DR PIR; A43001; A43001.
 DR PIR; P00012; P00012.
 DR InterPro; IPR001651; Gastrin.
 DR PROSITE; PS00259; GASTRIN; 1.
 KW Amidation; Sulfation; Hormone.
 FT MOD_RES 2 2 SULFATION.
 FT MOD_RES 8 8 AMIDATION.
 SQ SEQUENCE 8 AA; 1064 MW; DDCAA68378768B5A CRC64;

 Query Match 20.0%; Score 1; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 G 1
 DB 4 G 4

 RESULT 88
 CLP_THICU
 ID CLP_THICU STANDARD; PRT; 8 AA.
 AC P80488;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Chemolithotroph-specific protein (Fragment).
 OS Thiobacillus cuprinus.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Comamonadaceae; Thiomonas.
 OX NCBI_TaxID=36860;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=DSM 5494;
 RA Marin I., Amaro A.M., Jerez C.A., Amils R., Abad J.P.;
 RL Submitted (SEP-1995) to the SWISS-PROT data bank.
 CC -!- MISCELLANEOUS: FOUND SPECIFICALLY IN CELLS CULTURED
 CC CHEMOLITHOTROPHICALLY.
 FT NON_TER 8 8
 SQ SEQUENCE 8 AA; 785 MW; 91487B06DDC2D76D CRC64;

 Query Match 20.0%; Score 1; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 G 1
 DB 7 G 7

 RESULT 89
 COW2_CONPU
 ID COW2_CONPU STANDARD; PRT; 8 AA.
 AC P58785;

DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Leu-contryphan-P.
 OS Conus purpurascens (purple cone).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
 OC Apogastropoda; Caenogastropoda; Sorbeconcha; Hypogastropoda;
 OC Neogastropoda; Conoidea; Conidae; Conus.
 OX NCBI_TaxID=41690;
 RN [1]
 RP SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.
 RC STRAIN=Clipperton Island; TISSUE=Venom;
 RX MEDLINE=99388939; PubMed=10461743;
 RA Jacobsen R.B., Jimenez E.C., De la Cruz R.G.C., Gray W.R., Cruz L.J.,
 RA Oliveira B.M.;
 RT "A novel D-leucine-containing Conus peptide: diverse conformational
 RT dynamics in the contryphan family.";
 RL J. Pept. Res. 54:93-99(1999).
 CC -|- SUBCELLULAR LOCATION: Secreted.
 CC -|- TISSUE SPECIFICITY: Expressed by the venom duct.
 CC -|- MASS SPECTROMETRY: MW=888.4; METHOD=LSIMS.
 CC -|- SIMILARITY: BELONGS TO THE CONTRYPHAN FAMILY.
 CC Toxin; Hydroxylation; D-amino acid.
 KW DISULFID 2 8
 FT MOD RES 4 4 4 D-LEUCINE.
 FT SEQUENCE 8 AA; 890 MW; 75A367672732CEB8 CRC64;
 SQ
 Query Match 20.0%; Score 1; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 2y 1 G 1
 3b 1 G 1
 RESULT 90
 ID CPD1_ENTFA STANDARD; PRT; 8 AA.
 AC F13269;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-FEB-1991 (Rel. 17, Last annotation update)
 DE Sex pheromone cPDI.
 OS Enterococcus faecalis (Streptococcus faecalis).
 OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
 OX NCBI_TaxID=1351;
 RN [1]
 RP SEQUENCE.
 RC MEDLINE=85040388; PubMed=6436978;
 RA Suzuki A., Mori M., Sagakami Y., Isogai A., Fujino M., Kitada C.,
 RA Craig R.A., Clewell D.B.;
 RT "Isolation and structure of bacterial sex pheromone, cPDI.";
 RL Science 226:849-850(1984).
 CC -|- FUNCTION: cPDI IS INVOLVED IN THE CONJUGATIVE TRANSFER OF THE
 CC BACTERIOCIN PLASMID PPDI.
 CC Pheromone.
 CC SEQUENCE 8 AA; 913 MW; 8665B729C682C729 CRC64;
 SQ
 Query Match 20.0%; Score 1; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 2y 5 S 5
 3b 7 S 7
 RESULT 91
 ID FANRE STANDARD; PRT; 8 AA.
 AC P41872;
 DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE FMRFamide-like neuropeptide PFI (SDPNFLRF-amide).
 OS Panagrellus redivivus.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
 OC Panagrolaimoidea; Panagrolaimidae; Panagrellus.
 OX NCBI_TaxID=6233;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=93027659; PubMed=1408999;
 RA Geary T.G., Price D.A., Bowman J.W., Winterrowd C.A., Mackenzie C.D.,
 RA Garrison R.D., Williams J.F., Friedman A.R.;
 RT "Two FMRFamide-like peptides from the free-living nematode
 RT Panagrellus redivivus.";
 RL Peptides 13:209-214(1992).
 CC -|- FUNCTION: MYOACTIVE.
 CC -|- TISSUE SPECIFICITY: NERVE CORDS AND PAIRED GROUPS OF CELLS LOCATED
 CC CAUDALLY TO THE BASE OF THE PHARYNX.
 CC -|- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
 CC FAMILY.
 KW Neuropeptide; Amidation.
 FT MOD RES 8 8
 FT SEQUENCE 8 AA; 995 MW; C6D40729C4576AB5 CRC64;
 SQ
 Query Match 20.0%; Score 1; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 2y 5 S 5
 3b 1 S 1
 RESULT 92
 ID FAN1_PENMO STANDARD; PRT; 8 AA.
 AC P8316;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE FMRFamide-like neuropeptide FLPI (GDRNFLRF-amide).
 OS Penaeus monodon (Pencoid shrimp).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;
 OC Penaeidae; Penaeus.
 OX NCBI_TaxID=6687;
 RN [1]
 RP SEQUENCE, AND MASS SPECTROMETRY.
 RC TISSUE=Eyestalk;
 RX MEDLINE=21956277; PubMed=11959015;
 RA Sithigorngul P., Pupum J., Krungkarn S., Longyant S.,
 RA Chaivuthangkura P., Sithigorngul W., Petson A.;
 RT "Seven novel FMRFamide-like neuropeptide sequences from the eyestalk
 RT of the giant tiger prawn Penaeus monodon.";
 RL Comp. Biochem. Physiol. 131B:325-337(2002).
 CC -|- SUBCELLULAR LOCATION: Secreted.
 CC -|- MASS SPECTROMETRY: MW=1024.8; METHOD=MALDI.
 CC -|- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
 CC FAMILY.
 KW Neuropeptide; Amidation.
 FT MOD RES 8 8
 FT SEQUENCE 8 AA; 1034 MW; 72D40729C4540AAB CRC64;
 SQ
 Query Match 20.0%; Score 1; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 2y 1 G 1
 3b 1 G 1
 RESULT 93

```

FAR3_HOMAM
ID FAR3_HOMAM STANDARD; PRT; 8 AA.
AC P41486;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE FMRamide-like neuropeptide 3 (FLL 3) (F2).
OS Homarus americanus (American lobster).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
OC Nephropoidea; Nephropidae; Homarus.
OX NCBI_TaxID=6706;
RN [1]
RP SEQUENCE.
RC TISSUE=Pericardial organs;
RX MEDLINE=88116164; PubMed=3429714;
RA Trimmer B.A., Kobierski L.A., Kravitz E.A.;
RT "Purification and characterization of FMRamide-like immunoreactive
RT substances from the lobster nervous system: isolation and sequence
RT analysis of two closely related peptides.";
RL J. Comp. Neurol. 265:16-26(1987)
CC -!- MISCELLANEOUS: PERICARDIAL ORGANS RELEASE THIS PEPTIDE WITH 100 MM
CC -!- POTASSIUM IN THE PRESENCE OF CALCIUM.
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD RES 8 AMIDATION.
SQ SEQUENCE 8 AA; 1054 MW; C6D40729C4540AB5 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 5
DB 1 1

RESULT 94
FAR7_ASCSU
ID FAR7_ASCSU STANDARD; PRT; 8 AA.
AC P43171;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE FMRamide-like neuropeptide AF7.
OS Ascaris suum (Pig roundworm) (Ascaris lumbricoides).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC Ascarididae; Ascaris.
OX NCBI_TaxID=6253;
RN [1]
RP SEQUENCE.
RX MEDLINE=95380362; PubMed=7651904;
RA Cowden C., Stretton A.O.W.;
RT "Eight novel FMRamide-like neuropeptides isolated from the nematode
RT Ascaris suum.";
RL Peptides 15:491-500(1995).
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD RES 8 AMIDATION.
SQ SEQUENCE 8 AA; 963 MW; 9CD40059D417687D CRC64;

Query Match 20.0%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
DB 2 G 2

RESULT 95
FAR8_CALVO
ID FAR8_CALVO STANDARD; PRT; 8 AA.
AC P41863;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE CalliFMRamide 8.
OS Calliphora vomitoria (Blue blowfly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
OC Calliphoridae; Calliphora.
OX NCBI_TaxID=27454;
RN [1]
RP SEQUENCE.
RC TISSUE=Thoracic ganglion;
RX MEDLINE=92196111; PubMed=1549595;
RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,
RA Rehfeld J.F., Thorpe A.;
RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
RT neuropeptides (designated calliFMRamides) from the blowfly
RT Calliphora vomitoria";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992)
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD RES 8 AMIDATION.
SQ SEQUENCE 8 AA; 957 MW; 72D40699CAA44DD8 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
DB 1 G 1

RESULT 96
FUSS_FUSSO
ID FUSS_FUSSO STANDARD; PRT; 8 AA.
AC P81010;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 38, Last annotation update)
DE Allergen Fus s 13596* (Pragmat).
OS Fusarium solani (subsp. pisi) (Nectria haematococca).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Nectriaceae; Nectria.
OX NCBI_TaxID=70791;
RN [1]
RP SEQUENCE.
RC STRAIN=IARI 3596; TISSUE=Mycelium;
RA Verma J., Gangal S.V.;
RL Submitted (JUL-1997) to the SWISS-PROT data bank.
KW Allergen.
FT NON TER 8
SQ SEQUENCE 8 AA; 898 MW; C372C441F5B69041 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 S 5
DB 4 S 4

RESULT 97
GLUR_HUMAN
ID GLUR_HUMAN STANDARD; PRT; 8 AA.
AC P02729;
DT 21-JUL-1986 (Rel. 01, Created)

```


DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Urine glycopeptide.
DE Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RN SEQUENCE.
RP MEDLINE=72062338; PubMed=5126885;
RX Lote C.J., Weiss J.B.;
RA "Identification in urine of a low-molecular-weight highly polar
RA glycopeptide containing cysteinyl-galactose.";
RL Biochem. J. 123:25P-25P (1971).
RL -!- FUNCTION: THE IDENTITY OF THE GLYCOPROTEIN FROM WHICH THIS PEPTIDE
CC IS DERIVED IS UNKNOWN. NO PHYSIOLOGICAL FUNCTION HAS BEEN
CC ATTRIBUTED. AN ERYTHROCYTE MEMBRANE GLYCOPROTEIN HAVING A
CC SIMILAR STRUCTURE HAS ALSO BEEN FOUND.
CC PIR; A03188; XGHURU.
DR DR GO; GO:0005576; C:extracellular; NAS.
DR Glycoprotein. 1 1 S-LINKED (GAL. .).
KW CARBOHYD 1 1
FT SEQUENCE 8 AA; 855 MW; C2D87AA1F5B1E51E CRC64;
SQ
Query Match 20.0%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 S S
DB 4 S 4
RESULT 98
4TF1_PERAM
ID HTF1_PERAM STANDARD; PRT; 8 AA.
AC P04548;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypertrichalosemic factor 1 (Neuropeptide M-I) (Periplaneta CC-I)
DE (Pea-CAH-1) (Leu-CC-1) (Hypertrichalosemic neuropeptide 1).
DE Periplaneta americana (American cockroach),
DE Lepidoptera decemlineata (Colorado potato beetle), and
DE Blatta orientalis (Oriental cockroach).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blattodea;
OC Blattidae; Periplaneta.
OC NCBI_TaxID=6978, 7539, 6976;
RN [1]
RN SEQUENCE.
RP SPECIES=P.americana;
RC MEDLINE=85046530; PubMed=6548628;
RX Witten J.L., Schaffer M.H., O'Shea M., Cook J.C., Hemling M.E.,
RX Rinehart K.L. Jr.;
RT "Structures of two cockroach neuropeptides assigned by fast atom
RT bombardment mass spectrometry.";
RL Biochem. Biophys. Res. Commun. 124:350-358 (1984).
RN [2]
RN SEQUENCE.
RP SPECIES=P.americana;
RC MEDLINE=84298179; PubMed=6591205;
RX Scarborough R.M., Jamieson G.C., Kalish F., Kramer S.J., McEnroe G.A.,
RX Miller C.A., Schooley D.A.;
RA "Isolation and primary structure of two peptides with
RA cardioacceleratory and hyperglycemic activity from the corpora
RA cardiaca of Periplaneta americana";
RL Proc. Natl. Acad. Sci. U.S.A. 81:5575-5579 (1984).
RN [3]
RN SEQUENCE.
RP SPECIES=L.decemlineata; TISSUE=Corpora cardiaca;
RX MEDLINE=90160053; PubMed=2576128;
RX Gaede G., Kellner R.;

RT "The metabolic neuropeptides of the corpus cardiaca from the potato
RT beetle and the American cockroach are identical.";
RL Peptides 10:1287-1289 (1989).
RN [4]
RN SEQUENCE.
RP SPECIES=B.orientalis; TISSUE=Corpora cardiaca;
RC MEDLINE=90253659; PubMed=2340112;
RX Gaede G., Rinehart K.L. Jr.;
RT "Primary structures of hypertrichalosemic neuropeptides isolated from
RT the corpora cardiaca of the cockroaches Leucophaea maderae,
RT Gromphadorhina portentosa, Blattella germanica and Blatta orientalis
RT and of the stick insect Extatosoma tiaratum assigned by tandem fast
RT atom bombardment mass spectrometry.";
RL Biol. Chem. Hoppe-Seyler 371:345-354 (1990).
CC -!- FUNCTION: HYPERTRICHALOSEMIC FACTORS ARE NEUROPEPTIDES THAT
CC ELEVATE THE LEVEL OF TREHALOSE IN THE HEMOLYMPH (TREHALOSE IS
CC THE MAJOR CARBOHYDRATE IN THE HEMOLYMPH OF INSECTS).
CC -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR PIR; A44960; A44960.
DR PIR; A49823; A49823.
DR PIR; S08995; S08995.
DR InterPro; IPR002047; AKH.
DR PROSITE; PS00256; AKH; 1.
KW Neuropeptide; Amidation; Pyroglutamate carboxylic acid.
FT MOD_RES 1 1
FT MOD_RES 8 8 AMIDATION.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 991 MW; 86745775B9C452D6 CRC64;
Query Match 20.0%; Score 1; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 S S
DB 5 S 5
RESULT 99
HTF_TENMO
ID HTF_TENMO STANDARD; PRT; 8 AA.
AC P25419;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypertrichalosemic factor (HOTH) (Hypertrichalosemic neuropeptide).
DE Tenebrio molitor (Yellow mealworm), and
OS Zophobas rugipes.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
OC Tenebrionidae; Tenebrio.
OC NCBI_TaxID=7067, 7075;
RN [1]
RN SEQUENCE.
RP SPECIES=T.molitor, and Z.rugipes;
RC TISSUE=Corpora cardiaca;
RX MEDLINE=90341081; PubMed=2381871;
RX Gaede G., Rosinski G.;
RT "The primary structure of the hypertrichalosemic neuropeptide from
RT tenebrionid beetles: a novel member of the AKH/RPCH family.";
RL Peptides 11:455-459 (1990).
CC -!- FUNCTION: HYPERTRICHALOSEMIC FACTORS ARE NEUROPEPTIDES THAT
CC ELEVATE THE LEVEL OF TREHALOSE IN THE HEMOLYMPH (TREHALOSE IS
CC THE MAJOR CARBOHYDRATE IN THE HEMOLYMPH OF INSECTS).
CC -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR PIR; A43976; A43976.
DR PIR; B43976; B43976.
DR InterPro; IPR002047; AKH.
DR PROSITE; PS00256; AKH; 1.
KW Neuropeptide; Amidation; Pyroglutamate carboxylic acid.
FT MOD_RES 1 1
FT MOD_RES 8 8 AMIDATION.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 1005 MW; 86745775B9C44736 CRC64;

Query Match 20.0%; Score 1; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 S 5
 |
 Db 5 S 5

RESULT 100
 LCK1 LEUMA
 ID LCK1 LEUMA STANDARD; PRT; 8 AA.
 AC P21140;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 01-MAY-1991 (Rel. 18, Last annotation update)
 DE Leucokinin I (L-I).
 OS Leucophaea maderae (Madeira cockroach).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;
 OC Blaberidae; Leucophaea.
 OC NCBI_TaxID=6988;
 RN !1)
 RP SEQUENCE, AND SYNTHESIS.
 RC TISSUE=Head;
 RA Holman G.M., Cook B.J., Nachman R.J.;
 RT "Isolation, primary structure and synthesis of two neuropeptides
 RT from Leucophaea maderae: members of a new family of
 RT Cephalomyotropins.";
 RL Comp. Biochem. Physiol. 84C:205-211(1986).
 CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
 CC ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).
 CC -!- SIMILARITY: TO THE OTHER LEUCOKININS.
 KW Neuropeptide; Amidation.
 FT MOD RES 8
 SQ SEQUENCE 8 AA; 893 MW; DCG365B49CDC76A CRC64;

Query Match 20.0%; Score 1; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 S 5
 |
 Db 6 S 6

Search completed: October 21, 2003, 18:24:32
 Job time : 9.61111 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:24:06 ; Search time 16.1111 Seconds
(without alignments)

29.845 Million cell updates/sec

Title: US-10-057-890A-16

Perfect score: 5

Sequence: 1 GGGGS 5

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1100

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Listing first 150 summaries

Database : PIR 76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4	80.0	10	E60788	sperm-activating p
2	4	80.0	10	D60787	sperm-activating p
3	4	80.0	10	F60527	sperm-activating p
4	4	80.0	10	D60527	sperm-activating p
5	4	80.0	10	E60527	sperm-activating p
6	4	80.0	10	C60527	sperm-activating p
7	4	80.0	10	A60589	sperm-activating p
8	4	80.0	10	B60588	sperm-activating p
9	4	80.0	10	C60588	sperm-activating p
10	3	60.0	5	C53284	T-cell receptor be
11	3	60.0	6	JU0355	lipopeptide WS1279
12	3	60.0	7	E61491	seed protein ws-5
13	3	60.0	8	PT0509	T-cell receptor be
14	3	60.0	9	YFPG	thymic factor - pi
15	3	60.0	9	A60957	thymocyte growth p
16	3	60.0	10	F60787	sperm-activating p
17	3	60.0	10	H60787	sperm-activating p
18	3	60.0	10	C60788	sperm-activating p
19	3	60.0	10	D60788	sperm-activating p
20	3	60.0	10	C60787	sperm-activating p
21	3	60.0	10	A60787	sperm-activating p
22	3	60.0	10	B60787	sperm-activating p
23	3	60.0	10	A60788	sperm-activating p
24	3	60.0	10	A60527	sperm-activating p
25	3	60.0	10	G60527	sperm-activating p
26	3	60.0	10	C39572	sperm-activating p
27	3	60.0	10	E39572	sperm-activating p
28	3	60.0	10	F60589	sperm-activating p
29	3	60.0	10	D60589	sperm-activating p

30	3	60.0	10	2	C60589	sperm-activating p
31	3	60.0	10	2	B60589	sperm-activating p
32	3	60.0	10	2	I60527	sperm-activating p
33	3	60.0	10	2	A60588	sperm-activating p
34	3	60.0	10	2	D60588	sperm-activating p
35	3	60.0	10	2	PT0289	Ig heavy chain CRD
36	3	60.0	10	2	E41946	T-cell receptor ga
37	3	60.0	10	2	PH0925	T-cell receptor be
38	3	60.0	3	3	A23751	spinal cord peptid
39	2	40.0	4	2	T30569	hypothetical prote
40	2	40.0	4	2	A53284	T-cell receptor be
41	2	40.0	4	2	B53284	T-cell receptor be
42	2	40.0	4	2	PT0706	T-cell receptor be
43	2	40.0	4	2	S47552	ubiquitin - rat
44	2	40.0	5	2	A41225	copper resistance
45	2	40.0	5	2	B61445	Leu-enkephalin - b
46	2	40.0	5	2	A61445	Met-enkephalin - b
47	2	40.0	5	2	B61168	cocoonase (EC 3.4.
48	2	40.0	5	2	PT0267	Ig heavy chain CRD
49	2	40.0	5	2	D44823	synaptosomal-assoc
50	2	40.0	5	2	PT0659	T-cell receptor be
51	2	40.0	5	2	PT0707	T-cell receptor be
52	2	40.0	5	2	PT0585	T-cell receptor be
53	2	40.0	5	2	PT0717	T-cell receptor be
54	2	40.0	6	2	B61512	variant surface gl
55	2	40.0	6	2	A27696	contraction-inhibi
56	2	40.0	6	2	PT0280	Ig heavy chain CRD
57	2	40.0	6	2	PT0514	T-cell receptor be
58	2	40.0	6	2	PT0512	T-cell receptor be
59	2	40.0	6	2	PT0605	T-cell receptor be
60	2	40.0	6	2	PT0720	T-cell receptor be
61	2	40.0	6	2	PT0560	T-cell receptor be
62	2	40.0	6	2	PT0723	T-cell receptor be
63	2	40.0	6	2	PT0727	T-cell receptor be
64	2	40.0	6	2	PT0730	T-cell receptor be
65	2	40.0	6	2	A41946	T-cell receptor ga
66	2	40.0	7	2	A60224	Met-enkephalin-Arg
67	2	40.0	7	2	S42407	gramicidin S synth
68	2	40.0	7	2	H33098	180K exoantigen -
69	2	40.0	7	2	PT0523	T-cell receptor be
70	2	40.0	7	2	PT0642	T-cell receptor be
71	2	40.0	7	2	PT0667	T-cell receptor be
72	2	40.0	7	2	PT0666	T-cell receptor be
73	2	40.0	7	2	PT0663	T-cell receptor be
74	2	40.0	7	2	PT0542	T-cell receptor be
75	2	40.0	7	2	PT0683	T-cell receptor be
76	2	40.0	7	2	PT0719	T-cell receptor be
77	2	40.0	7	2	PT0586	T-cell receptor be
78	2	40.0	7	2	PT0702	T-cell receptor be
79	2	40.0	7	2	PT0728	T-cell receptor be
80	2	40.0	7	2	A38671	peptidylglycine mo
81	2	40.0	7	2	S58797	serine/threonine-s
82	2	40.0	7	2	A58718	carnocin U149 - Ca
83	2	40.0	7	4	PC2056	trichodecenen I -
84	2	40.0	7	4	PC2057	trichodecenen II -
85	2	40.0	7	4	SI5597	orf 4 rara 5'-regi
86	2	40.0	8	2	PL0184	capsid protein VP-
87	2	40.0	8	2	A21440	variant surface gl
88	2	40.0	8	2	JS0315	leucokinin V - Mad
89	2	40.0	8	2	A61467	penalbin - Adeli
90	2	40.0	8	2	PT0279	Ig heavy chain CRD
91	2	40.0	8	2	PT0323	Ig heavy chain CRD
92	2	40.0	8	2	I48934	apolipoprotein A-1
93	2	40.0	8	2	A35768	T-cell receptor al
94	2	40.0	8	2	PT0595	T-cell receptor be
95	2	40.0	8	2	PT0530	T-cell receptor be
96	2	40.0	8	2	PT0527	T-cell receptor be
97	2	40.0	8	2	PT0639	T-cell receptor be
98	2	40.0	8	2	PT0557	T-cell receptor be
99	2	40.0	8	2	PT0716	T-cell receptor be
100	2	40.0	8	2	PT0725	T-cell receptor be
101	2	40.0	8	2	PT0588	T-cell receptor be
102	2	40.0	8	2	PC4372	telomeric and tetr

103 2 40.0 8 4 I55411
 104 2 40.0 9 2 A44873
 105 2 40.0 9 2 F28854
 106 2 40.0 9 2 D24180
 107 2 40.0 9 2 A60522
 108 2 40.0 9 2 QDRB
 109 2 40.0 9 2 B45020
 110 2 40.0 9 2 A13576
 111 2 40.0 9 2 T31612
 112 2 40.0 9 2 A57444
 113 2 40.0 9 2 B57444
 114 2 40.0 9 2 C57444
 115 2 40.0 9 2 D57444
 116 2 40.0 9 2 A53797
 117 2 40.0 9 2 I50633
 118 2 40.0 9 2 JQ0914
 119 2 40.0 9 2 PT0225
 120 2 40.0 9 2 A37027
 121 2 40.0 9 2 PC2197
 122 2 40.0 9 2 S65865
 123 2 40.0 9 2 FH1591
 124 2 40.0 9 2 S36850
 125 2 40.0 9 2 B38740
 126 2 40.0 9 2 PT0670
 127 2 40.0 9 2 PT0634
 128 2 40.0 9 2 PT0562
 129 2 40.0 9 2 G41946
 130 2 40.0 10 2 S24190
 131 2 40.0 10 2 A61131
 132 2 40.0 10 2 A60410
 133 2 40.0 10 2 PQ0177
 134 2 40.0 10 2 PC2172
 135 2 40.0 10 2 G60787
 136 2 40.0 10 2 E60787
 137 2 40.0 10 2 E60589
 138 2 40.0 10 2 I60588
 139 2 40.0 10 2 H60588
 140 2 40.0 10 2 S06964
 141 2 40.0 10 2 S70336
 142 2 40.0 10 2 A27617
 143 2 40.0 10 2 C61440
 144 2 40.0 10 2 D61440
 145 2 40.0 10 2 PN0165
 146 2 40.0 10 2 B61512
 147 2 40.0 10 2 B56899
 148 2 40.0 10 2 A24196
 149 2 40.0 10 2 S43630
 150 2 40.0 10 2 PT0245

ALIGNMENTS

RESULT 1
 E60788
 sperm-activating peptide (Ala-3, Gly-5 speract) - sea urchin (Pseudocentrotus depressus)
 C:Species: Pseudocentrotus depressus
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 18-Aug-2000
 C:Accession: E60788
 R;Yoshino, K.; Kajiwara, H.; Nomura, K.; Garbers, D.L.; Yoshino, K.; Kurita, M.; Tanaka, H.
 Comp. Biochem. Physiol. B 89, 687-693, 1988
 A;Title: Some more speract derivatives associated with eggs of sea urchins, Pseudocentrotus
 A;Reference number: A60787; MUID:88242184; PMID:3378407
 A;Accession: E60788
 A;Molecule type: protein
 A;Residues: 1-10 <SUZ>
 C:Comment: This oligopeptide from egg jelly is one of several from this species, all of
 at shows some, but not absolute, species restriction.
 C:Superfamily: unassigned animal peptides

Query Match 80.0%; Score 4; DB 2; Length 10;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGG 4
 DB 5 GGGG 8

RESULT 2

D60787
 sperm-activating peptide (Ser-1, Ala-3, Gly-5 speract) - sea urchin (Hemicentrotus pulch-
 C:Species: Hemicentrotus pulcherrimus
 C:Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 18-Aug-2000
 C:Accession: D60787
 R;Suzuki, N.; Kajiwara, H.; Nomura, K.; Garbers, D.L.; Yoshino, K.; Kurita, M.; Tanaka, H.
 Comp. Biochem. Physiol. B 89, 687-693, 1988
 A;Title: Some more speract derivatives associated with eggs of sea urchins, Pseudocentrotus
 A;Reference number: A60787; MUID:88242184; PMID:3378407
 A;Accession: D60787
 A;Molecule type: protein
 A;Residues: 1-10 <SUZ>
 C:Comment: This oligopeptide from egg jelly is one of several from this species, all of
 at shows some, but not absolute, species restriction.
 C:Superfamily: unassigned animal peptides

Query Match 80.0%; Score 4; DB 2; Length 10;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGG 4
 DB 5 GGGG 8

RESULT 3

F60527
 sperm-activating peptide (Gly-3,5 SAP-I) - sea urchin (Tripneustes gratilla)
 N;Alternate names: speract homolog TG-8
 C:Species: Tripneustes gratilla
 C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 18-Aug-2000
 C:Accession: F60527; H39572
 R;Yoshino, K.; Kajiwara, H.; Nomura, K.; Takao, T.; Shimonishi, Y.; Kurita, M.; Yamaguchi
 Comp. Biochem. Physiol. B 94, 739-751, 1989
 A;Title: A halogenated amino acid-containing sperm activating peptide and its related pe
 otus nudus, Echinometra mathaei and Heterocentrotus mamillatus.
 A;Reference number: A60527
 A;Accession: F60527
 A;Molecule type: protein
 A;Residues: 1-10 <YOS>
 A;Note: this peptide did not contain bromophenylalanine
 R;Yoshino, K.; Takao, T.; Suhara, M.; Kitai, T.; Hori, H.; Nomura, K.; Yamaguchi, M.; Sh
 Biochemistry 30, 6203-6209, 1991
 A;Title: Identification of a novel amino acid, o-bromo-L-phenylalanine, in egg-associate
 A;Reference number: A39572; MUID:91283461; PMID:2059627

Query Match 80.0%; Score 4; DB 2; Length 10;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGG 4
 DB 5 GGGG 8

RESULT 4

D60527
 sperm-activating peptide (Gly-5 SAP-I) - sea urchin (Tripneustes gratilla)
 N;Alternate names: speract homolog TG-6

N;Alternate names: speract homolog TG-4
C;Species: Tripneustes gratilla
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 18-Aug-2000
C;Accession: C60527; D39572
R;Yoshino, K.I.; Kajiuura, H.; Nomura, K.; Takao, T.; Shimonishi, Y.; Kurita, M.; Yama
Comp. Biochem. Physiol. B 94, 739-751, 1989
A;Title: A halogenated amino acid-containing sperm activating peptide and its related
otus nudus, Echinometra mathaei and Heterocentrotus mammillatus.
A;Reference number: A60527
A;Accession: C60527
A;Molecule type: protein
A;Residues: 1-10 <YOS>
A;Note: this peptide did not contain bromophenylalanine
R;Yoshino, K.; Takao, T.; Suhara, M.; Kitai, T.; Hori, H.; Nomura, K.; Yamaguchi, M.;
Biochemistry 30, 6203-6209, 1991
A;Title: Identification of a novel amino acid, o-bromo-L-phenylalanine, in egg-associ
A;Reference number: A39572; MUID:91283461; PMID:2059627
A;Accession: D39572
A;Molecule type: protein
A;Residues: 1-10 <Y02>
A;Note: this peptide contained bromophenylalanine
C;Superfamily: unassigned animal peptides
C;Keywords: bromine
F;2'Modified site: 2'-bromophenylalanine (Phe) (partial) #status experimental

Query Match	80.0%;	Score 4;	DB 2;	Length 10;
Best Local Similarity	100.0%;	Pred. No. 2.3e+02;		
Matches	4;	Conservative 0;	Mismatches 0;	Indels 0;
Gy	1	GGGG 4		
Dd	5	GGGG 8		

RESULT 7

A60589
sperm-activating peptide (Tyr-2, Asn-3, Gly-5, Ile-9, Asp-10 SAP-I) - slate-pencil ur
C;Species: Heterocentrotus mammillatus
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 18-Aug-2000
C;Accession: A60589
R;Yoshino, K.I.; Kajiuura, H.; Nomura, K.; Takao, T.; Shimonishi, Y.; Kurita, M.; Yama
Comp. Biochem. Physiol. B 94, 739-751, 1989
A;Title: A halogenated amino acid-containing sperm activating peptide and its related
otus nudus, Echinometra mathaei and Heterocentrotus mammillatus.
A;Reference number: A60527
A;Accession: A60589
A;Molecule type: protein
A;Residues: 1-10 <YOS>
C;Superfamily: unassigned animal peptides

Query Match	80.0%;	Score 4;	DB 2;	Length 10;
Best Local Similarity	100.0%;	Pred. No. 2.3e+02;		
Matches	4;	Conservative 0;	Mismatches 0;	Indels 0;
Gy	1	GGGG 4		
Dd	5	GGGG 8		

RESULT 8

B60588
sperm-activating peptide (Ala-3, Gly-5 SAP-I) - sea urchin (Strongylocentrotus nudus)
N;Alternate names: speract homolog
C;Species: Strongylocentrotus nudus
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 18-Aug-2000
C;Accession: B60588
R;Yoshino, K.I.; Kajiuura, H.; Nomura, K.; Takao, T.; Shimonishi, Y.; Kurita, M.; Yama
Comp. Biochem. Physiol. B 94, 739-751, 1989
A;Title: A halogenated amino acid-containing sperm activating peptide and its related
otus nudus, Echinometra mathaei and Heterocentrotus mammillatus.
A;Reference number: A60527
A;Accession: B60588
A;Molecule type: protein

Okada, Y.; Tsuda, Y.; Tanaka, M.; Shigematsu, N.; Hori, Y.; Goto, T.; Hashimoto, M.
J. Pharm. Bull. 39, 607-611, 1991

N:Alternate names: FTS (facteur thymique serique)
 C:Species: Sus scrofa domestica (domestic pig)
 C>Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 07-May-1999
 C:Accession: A01523; A60983
 R:Pleau, J.M.; Dardenne, M.; Blouquit, Y.; Bach, J.F.
 J. Biol. Chem. 252, 8045-8047, 1977
 A:Title: Structural study of circulating thymic factor: a peptide isolated from pig serum
 A:Reference number: A01523; MUID:78026571; PMID:914862
 A:Accession: A01523
 A:Molecule type: protein
 A:Residues: 1-9 <PLE>
 R:Bach, J.F.; Dardenne, M.; Pleau, J.M.; Rosa, J.
 Nature 266, 55-57, 1977
 A:Title: Biochemical characterisation of a serum thymic factor.
 A:Reference number: A60983; MUID:77123829; PMID:300146
 A:Accession: A60983
 A:Molecule type: protein
 A:Residues: 'Z', '2', '4', 'Z', '6'-9 <BAC>
 C:Comment: This peptide induces DNA synthesis in immature thymocytes, but not peripheral in a variety of immunosays.
 C:Comment: See PIR:A60957 (sheep) for discussion of another possible N-terminal modification
 C:Superfamily: thymic factor
 C:Keywords: pyrrolutamic acid
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 60.0%; Score 3; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05; Indels 0; Gaps 0;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 GGS 5
 DB 6 GGS 8

RESULT 15
 A60957
 thymocyte growth peptide - sheep
 N:Contains: FTS (facteur thymique serique)
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
 C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 24-Nov-1999
 C:Accession: A60957
 R:Ernstroem, U.; Gafvelin, G.; Rudja, J.M.
 Bioeci. Rep. 10, 403-412, 1990
 A:Title: Purification of thymocyte growth peptide (TGP) from sheep thymus. Relationship
 A:Reference number: A60957; MUID:91064427; PMID:2249004
 A:Accession: A60957
 A:Molecule type: protein
 A:Residues: 1-9 <ERN>
 C:Comment: This peptide induces DNA synthesis in immature thymocytes, but not peripheral in a variety of immunosays.
 C:Comment: This peptide was isolated in two forms. One form contained the pyrrolidone carboxylic acid (Gln) at the N-terminus, the other form contained the pyrrolidone carboxylic acid (Glu) at the N-terminus.
 C:Superfamily: thymic factor
 C:Keywords: blocked amino end; pyrrolutamic acid
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) (in thymocyte growth peptide) #status experimental
 F:1/Modified site: blocked amino end (Gln) (in thymocyte growth peptide) #status experimental

Query Match 60.0%; Score 3; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05; Indels 0; Gaps 0;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 GGS 5
 DB 6 GGS 8

RESULT 16
 F60787
 sperm-activating peptide (Ser-5 speract) - sea urchin (Anthocidaris crassispina)
 C:Species: Anthocidaris crassispina
 C>Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 18-Aug-2000
 C:Accession: F60787
 R:Suzuki, N.; Kajiwara, H.; Nomura, K.; Garbers, D.L.; Yoshino, K.; Kurita, M.; Tanaka, H.

Comp. Biochem. Physiol. B 89, 687-693, 1988
 A:Title: Some more speract derivatives associated with eggs of sea urchins, Pseudocentrotus
 A:Reference number: A60787; MUID:88242184; PMID:3378407
 A:Accession: F60787
 A:Molecule type: protein
 A:Residues: 1-10 <SUZ>
 C:Comment: This oligopeptide from egg jelly is one of several from this species, all of which show some, but not absolute, species restriction.
 C:Superfamily: unassigned animal peptides

Query Match 60.0%; Score 3; DB 2; Length 10;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03; Indels 0; Gaps 0;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGS 3
 DB 6 GGS 8

RESULT 17
 H60787
 sperm-activating peptide (Thr-5 speract) - sea urchin (Anthocidaris crassispina)
 C:Species: Anthocidaris crassispina
 C>Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 18-Aug-2000
 C:Accession: H60787
 R:Suzuki, N.; Kajiwara, H.; Nomura, K.; Garbers, D.L.; Yoshino, K.; Kurita, M.; Tanaka, H.
 Comp. Biochem. Physiol. B 89, 687-693, 1988
 A:Title: Some more speract derivatives associated with eggs of sea urchins, Pseudocentrotus
 A:Reference number: A60787; MUID:88242184; PMID:3378407
 A:Accession: H60787
 A:Molecule type: protein
 A:Residues: 1-10 <SUZ>
 C:Comment: This oligopeptide from egg jelly is one of several from this species, all of which show some, but not absolute, species restriction.
 C:Superfamily: unassigned animal peptides

Query Match 60.0%; Score 3; DB 2; Length 10;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03; Indels 0; Gaps 0;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGS 3
 DB 6 GGS 8

RESULT 18
 C60788
 sperm-activating peptide (speract) - sea urchin (Pseudocentrotus depressus)
 C:Species: Pseudocentrotus depressus
 C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 18-Aug-2000
 C:Accession: C60788
 R:Suzuki, N.; Kajiwara, H.; Nomura, K.; Garbers, D.L.; Yoshino, K.; Kurita, M.; Tanaka, H.
 Comp. Biochem. Physiol. B 89, 687-693, 1988
 A:Title: Some more speract derivatives associated with eggs of sea urchins, Pseudocentrotus
 A:Reference number: A60787; MUID:88242184; PMID:3378407
 A:Accession: C60788
 A:Molecule type: protein
 A:Residues: 1-10 <SUZ>
 C:Comment: This oligopeptide from egg jelly is one of several from this species, all of which show some, but not absolute, species restriction.
 C:Superfamily: unassigned animal peptides

Query Match 60.0%; Score 3; DB 2; Length 10;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03; Indels 0; Gaps 0;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGS 3
 DB 6 GGS 8

RESULT 19
 D60788

```

perm-activating peptide (Thr-5 speract) - sea urchin (Pseudocentrotus depressus)
Species: Pseudocentrotus depressus
Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 18-Aug-2000
Accession: D60788
Suzuki, N.; Kajiuura, H.; Nomura, K.; Garbers, D.L.; Yoshino, K.; Kurita, M.; Tanaka, H.
omp. Biochem. Physiol. B 89, 687-693, 1988
Title: Some more speract derivatives associated with eggs of sea urchins, Pseudocentrotus
Reference number: A60787; MUID:88242184; PMID:3378407
Accession: D60788
Molecule type: protein
Residues: 1-10 <SUZ>
Comment: This oligopeptide from egg jelly is one of several from this species, all of
t shows some, but not absolute, species restriction.
Superfamily: unassigned animal peptides

Query Match 60.0%; Score 3; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGG 3
b 6 GGG 8

RESULT 20
perm-activating peptide (Ser-3,10 speract) - sea urchin (Hemicentrotus pulcherrimus)
Species: Hemicentrotus pulcherrimus
Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 18-Aug-2000
Accession: C60787
Suzuki, N.; Kajiuura, H.; Nomura, K.; Garbers, D.L.; Yoshino, K.; Kurita, M.; Tanaka, H.
omp. Biochem. Physiol. B 89, 687-693, 1988
Title: Some more speract derivatives associated with eggs of sea urchins, Pseudocentrotus
Reference number: A60787; MUID:88242184; PMID:3378407
Accession: C60787
Molecule type: protein
Residues: 1-10 <SUZ>
Comment: This oligopeptide from egg jelly is one of several from this species, all of
t shows some, but not absolute, species restriction.
Superfamily: unassigned animal peptides

Query Match 60.0%; Score 3; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGG 3
b 6 GGG 8

RESULT 21
perm-activating peptide (speract) - sea urchin (Hemicentrotus pulcherrimus)
Species: Hemicentrotus pulcherrimus
Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 18-Aug-2000
Accession: A60787
Suzuki, N.; Kajiuura, H.; Nomura, K.; Garbers, D.L.; Yoshino, K.; Kurita, M.; Tanaka, H.
omp. Biochem. Physiol. B 89, 687-693, 1988
Title: Some more speract derivatives associated with eggs of sea urchins, Pseudocentrotus
Reference number: A60787; MUID:88242184; PMID:3378407
Accession: A60787
Molecule type: protein
Residues: 1-10 <SUZ>
Comment: This oligopeptide from egg jelly is one of several from this species, all of
t shows some, but not absolute, species restriction.
Superfamily: unassigned animal peptides

Query Match 60.0%; Score 3; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGG 3
b 6 GGG 8

RESULT 22
perm-activating peptide (Thr-5 speract) - sea urchin (Hemicentrotus pulcherrimus)
Species: Hemicentrotus pulcherrimus
Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 18-Aug-2000
Accession: B60787
R; Suzuki, N.; Kajiuura, H.; Nomura, K.; Garbers, D.L.; Yoshino, K.; Kurita, M.; Tanaka, H.
omp. Biochem. Physiol. B 89, 687-693, 1988
Title: Some more speract derivatives associated with eggs of sea urchins, Pseudocentrotus
Reference number: A60787; MUID:88242184; PMID:3378407
Accession: B60787
Molecule type: protein
Residues: 1-10 <SUZ>
Comment: This oligopeptide from egg jelly is one of several from this species, all of
t shows some, but not absolute, species restriction.
Superfamily: unassigned animal peptides

Query Match 60.0%; Score 3; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGG 3
b 6 GGG 8

RESULT 23
perm-activating peptide (Thr-5 speract) - sea urchin (Strongylocentrotus purpuratus);
Species: Strongylocentrotus purpuratus (purple urchin)
Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 18-Aug-2000
Accession: A60788
R; Suzuki, N.; Kajiuura, H.; Nomura, K.; Garbers, D.L.; Yoshino, K.; Kurita, M.; Tanaka, H.
omp. Biochem. Physiol. B 89, 687-693, 1988
Title: Some more speract derivatives associated with eggs of sea urchins, Pseudocentrotus
Reference number: A60787; MUID:88242184; PMID:3378407
Accession: A60788
Molecule type: protein
Residues: 1-10 <SUZ>
Comment: This oligopeptide from egg jelly is one of several from this species. Unlike
t shows some, but not absolute, species restriction (see PIR:A34543).
Superfamily: unassigned animal peptides

Query Match 60.0%; Score 3; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGG 3
b 6 GGG 8

RESULT 24
perm-activating peptide (Br-Phe-2 SAP-I) - sea urchin (Tripneustes gratilla)
N; Alternate names: SAP-a; speract; TG-1; TG-9
Species: Tripneustes gratilla
Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 18-Aug-2000
Accession: A60527; A39572; A60973; H60527; I39572
R; Yoshino, K.; Kajiuura, H.; Nomura, K.; Takao, T.; Shimonishi, Y.; Kurita, M.; Yamagu.
omp. Biochem. Physiol. B 94, 739-751, 1989
Title: A halogenated amino acid-containing sperm activating peptide and its related p
otus nudus, Echinometra mathaei and Heterocentrotus mammillatus.
Reference number: A60527
Accession: A60527
Molecule type: protein
Residues: 1-10 <YOS>
R; Yoshino, K.; Takao, T.; Suhara, M.; Kitai, T.; Hori, H.; Nomura, K.; Yamaguchi, M.; S
Biochemistry 30, 6203-6209, 1991
Title: Identification of a novel amino acid, o-bromo-L-phenylalanine, in egg-associat

```


A:Molecule type: protein
A:Residues: 1-10 <YOS>
C:Superfamily: unassigned animal peptides
C:Keywords: bromine
F:2/Modified site: 2'-bromophenylalanine (Phe) #status experimental

Query Match 60.0%; Score 3; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGG 3
|||
Db 6 GGG 8

RESULT 27
E39572
sperm-activating peptide TG-5 - sea urchin (Tripneustes gratilla)
N:Alternate names: speract homolog TG-5
C:Species: Tripneustes gratilla
C:date: 10-Jul-1992 #sequence_revision 10-Jul-1992 #text_change 18-Aug-2000
C:Accession: E39572
R:Yoshino, K.; Takao, T.; Suhara, M.; Kitai, T.; Hori, H.; Nomura, K.; Yamaguchi, M.;
Biochemistry 30, 6203-6209, 1991
A:title: Identification of a novel amino acid, o-bromo-L-phenylalanine, in egg-associated
A:reference number: A39572; MUID:91283461; PMID:2059627
A:Accession: E39572
A:Molecule type: protein
A:Residues: 1-10 <YOS>
C:Superfamily: unassigned animal peptides
C:Keywords: bromine
F:2/Modified site: 2'-bromophenylalanine (Phe) #status experimental

Query Match 60.0%; Score 3; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGG 3
|||
Db 6 GGG 8

RESULT 28
F60589
sperm-activating peptide (Asn-3, Ser-5 SAP-I) - Echinometra mathaei (type A)
C:Species: Echinometra mathaei
C:date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 18-Aug-2000
C:Accession: F60589
R:Yoshino, K.I.; Kajiyura, H.; Nomura, K.; Takao, T.; Shimonishi, Y.; Kurita, M.; Yamai
Comp. Biochem. Physiol. B 94, 739-751, 1989
A:Title: A halogenated amino acid-containing sperm activating peptide and its related
otus nudus, Echinometra mathaei and Heterocentrotus mammillatus.
A:reference number: A60527
A:Accession: F60589
A:Molecule type: protein
A:Residues: 1-10 <YOS>
C:Superfamily: unassigned animal peptides

Query Match 60.0%; Score 3; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGG 3
|||
Db 6 GGG 8

RESULT 29
D60589
sperm-activating peptide (Ser-3,5 SAP-I) - Echinometra mathaei
C:Species: Echinometra mathaei
C:date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 18-Aug-2000
C:Accession: D60589

1 **GGG 3**

Db 6 GGG 8

RESULT 35

PT0289
IG heavy chain CRD3 region (clone 4-109) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C:Accession: PT0289
R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and J
A:Reference number: PT0222; MUID:91108337; PMID:1899102
A:Accession: PT0289
A:Molecule type: DNA
A:Residues: 1-10 <YAM>
A:Experimental source: B lymphocyte
C:Keywords: heterotetramer; immunoglobulin

Query Match 60.0%; Score 3; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.7e+03; Indels 0; Gaps 0;
Matches 3; Conservative 0; Mismatches 0;

Qy 1 GGG 3
|||
Db 8 GGG 10

RESULT 36

E41946
T-cell receptor gamma chain (1a.9) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
C:Accession: E41946
R:Whetzel, M.; Mosley, R.L.; Whetzel, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.
Mol. Cell. Biol. 11, 5902-5909, 1991
A:Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma ge
A:Reference number: A41946; MUID:92049316; PMID:1658619
A:Accession: E41946
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-10 <WHE>
C:Keywords: T-cell receptor

Query Match 60.0%; Score 3; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.7e+03; Indels 0; Gaps 0;
Matches 3; Conservative 0; Mismatches 0;

Qy 1 GGG 3
|||
Db 7 GGG 9

RESULT 37

PH0925
T-cell receptor beta chain V-D-J region (isolate 11) - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 30-May-1997
C:Accession: PH0925
R:Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenberg, A.A.; Wilson, D.B.
J. Exp. Med. 174, 1467-1476, 1991
A:Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergi
A:Reference number: PH0891; MUID:92078857; PMID:1836012
A:Accession: PH0925
A:Molecule type: mRNA
A:Residues: 1-10 <GOL>
A:Experimental source: concanavalin A-activated lymphoblast
C:Keywords: T-cell receptor

Query Match 60.0%; Score 3; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.7e+03; Indels 0; Gaps 0;
Matches 3; Conservative 0; Mismatches 0;

Qy 1 GGG 3
|||
Db 6 GGG 8

RESULT 38

A23751
spinal cord peptide SCP-4 - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: A23751
R:Hei, K.L.; Chen, R.L.; Chen, Z.G.; Zhang, H.L.; Lu, Y.A.; Guo, S.Y.; Wu, S.X.; Tsou
Arch. Biochem. Biophys. 240, 178-183, 1985
A:Reference number: A23751; MUID:85250425; PMID:4015098
A:Accession: A23751
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-3 <HSI>
C:Superfamily: unassigned animal peptides

Query Match 40.0%; Score 2; DB 3; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.8e+05; Indels 0; Gaps 0;
Matches 2; Conservative 0; Mismatches 0;

Qy 1 GG 2
||
Db 2 GG 3

RESULT 39

T30569
hypothetical protein - Emericella nidulans
C:Species: Emericella nidulans, Aspergillus nidulans
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
C:Accession: T30569
R:Morrice, J.; Mackenzie, D.A.; Parr, A.J.; Archer, D.B.
Curr. Genet. 34, 379-385, 1998
A:Title: Isolation and characterisation of the acetyl-CoA carboxylase gene from Asper
A:Reference number: Z20869; MUID:99087906; PMID:9871120
A:Accession: T30569
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-4 <MOR>
A:Cross-references: EMBL:X15996; NID:e1285512; PID:e1218041; PIDN:CAA75927.1

Query Match 40.0%; Score 2; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05; Indels 0; Gaps 0;
Matches 2; Conservative 0; Mismatches 0;

Qy 4 GS 5
||
Db 2 GS 3

RESULT 40

A53284
T-cell receptor beta 2 chain D region, Dbeta2 - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 02-May-1994 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C:Accession: A53284
R:Harindranath, N.; Alexander, C.B.; Mage, R.G.
Mol. Immunol. 28, 881-888, 1991
A:Title: Evolutionarily conserved organization and sequences of germline diversity an
A:Reference number: A53284; MUID:91342695; PMID:1678859
A:Accession: A53284
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-4 <HAG>
A:Cross-references: GB:S60737; NID:g233916; PIDN:AA819517.1; PID:g233917
A:Note: sequence extracted from NCBI backbone (NCBIN:60737; NCBI:P:60739)
C:Keywords: T-cell receptor

Query Match 40.0%; Score 2; DB 2; Length 4;

Best Local Similarity 100.0%; Pred. No. 2.8e+05; Mismatches 0; Indels 0; Gaps 0;
Matches 2; Conservative 0;

Y 1 GG 2
||
b 3 GG 4

RESULT 41

cell receptor beta 2 chain D region, Dbeta2 - rabbit
Species: Oryctolagus cuniculus (domestic rabbit)
Date: 02-May-1994 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
Accession: B53284
Harindranath, N.; Alexander, C.B.; Mage, R.G.
J. Biol. Immunol. 28, 881-888, 1991
Title: Evolutionarily conserved organization and sequences of germline diversity and J
Reference number: A53284; MUID:91342695; PMID:1678859
Accession: B53284
Status: preliminary
Molecule type: DNA
Residues: 1-4 <HAR>
Cross-references: GB:S60737; NID:G233916; PID:AA819518.1; PID:G233918
Note: sequence extracted from NCBI backbone (NCBIN:60737, NCBIP:60738)
Keywords: T-cell receptor

Query Match 40.0%; Score 2; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GG 2
||
b 3 GG 4

RESULT 42

cell receptor beta chain V-D-J region (135-1DU) - mouse (fragment)
Species: Mus musculus (house mouse)
Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
Accession: PT0706
Feeney, A.J.
Exp. Med. 174, 115-124, 1991
Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
Reference number: PT0509; MUID:91277601; PMID:1711558
Accession: PT0706
Status: translation not shown
Molecule type: DNA
Residues: 1-4 <FE>
Experimental source: newborn thymus, strain BALB/c
Keywords: T-cell receptor

Query Match 40.0%; Score 2; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GG 2
||
b 3 GG 4

RESULT 43

biquitin - rat
Species: Rattus norvegicus (Norway rat)
Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 17-Mar-1999
Accession: S47552
Hubbard, M.J.; Carne, A.
Biochim. Biophys. Acta 1200, 191-196, 1994
Title: Differential feeding-related regulation of ubiquitin and calbindin(9kDa) in rat
Reference number: S47552; MUID:94304928; PMID:8031840
Accession: S47552
Status: preliminary

A,Molecule type: protein
A,Residues: 1-4 <HUB>

Query Match 40.0%; Score 2; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
||
DB 3 GG 4

RESULT 44

A41225
copper resistance protein A - Pseudomonas syringae pv. tomato (fragment)
C,Species: Pseudomonas syringae pv. tomato
C,Date: 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change 24-Jun-1993
C,Accession: A41225
R,Cha, J.S.; Cooksey, D.A.
Proc. Natl. Acad. Sci. U.S.A. 88, 8915-8919, 1991
A,Title: Copper resistance in Pseudomonas syringae mediated by periplasmic and outer mem
A,Reference number: A41225; MUID:92020961; PMID:1924351
A,Accession: A41225
A,Status: preliminary
A,Molecule type: protein
A,Residues: 1-5 <CHA>

Query Match 40.0%; Score 2; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
||
DB 3 GS 4

RESULT 45

B61445
Leu-enkephalin - blue mussel
C,Species: Mytilus edulis (blue mussel)
C,Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 21-Jan-2000
C,Accession: B61445
R,Leung, M.K.; Stefano, G.B.
Proc. Natl. Acad. Sci. U.S.A. 81, 955-958, 1984
A,Title: Isolation and identification of enkephalins in pedal ganglia of Mytilus edulis
A,Reference number: A61445; MUID:84144823; PMID:6583690
A,Accession: B61445
A,Molecule type: protein
A,Residues: 1-5 <LEU>
A,Experimental source: pedal ganglia
C,Keywords: neuropeptide; opioid peptide

Query Match 40.0%; Score 2; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
||
DB 2 GG 3

RESULT 46

A61445
Met-enkephalin - blue mussel
C,Species: Mytilus edulis (blue mussel)
C,Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 21-Jan-2000
C,Accession: A61445
R,Leung, M.K.; Stefano, G.B.
Proc. Natl. Acad. Sci. U.S.A. 81, 955-958, 1984
A,Title: Isolation and identification of enkephalins in pedal garglia of Mytilus edulis
A,Reference number: A61445; MUID:84144823; PMID:6583690
A,Accession: A61445
A,Molecule type: protein

A:Residues: 1-5 <LEU>
 A:Experimental source: pedal ganglia
 C:Keywords: neuropeptide; opioid peptide

Query Match 40.0%; Score 2; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
 ||
 Db 2 GG 3

RESULT 47

B61168

cocoonase (EC 3.4.21.-) - Chinese oak silkmoth (fragment)

C:Species: Antherea pernyi (Chinese oak silkmoth)

C:Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 07-May-1999

C:Accession: B61168

R:Kramer, K.J.; Felsted, R.L.; Law, J.H.

J. Biol. Chem. 248, 3021-3028, 1973

A:Title: Cocoonase. V. Structural studies on an insect serine protease.

A:Reference number: A61168; MUID:73166540; PMID:4735570

C:Accession: B61168

A:Molecule type: protein

A:Residues: 1-5 <KRA>

C:Keywords: hydrolase; serine proteinase; zymogen

F/1-5/Product: cocoonase (fragment) #status experimental <MAT>

Query Match 40.0%; Score 2; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
 ||
 Db 3 GG 4

RESULT 48

PT0267

Ig heavy chain CDR3 region (clone 3-94A) - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996

C:Accession: PT0267

R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.

J. Exp. Med. 173, 395-407, 1991

A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and

A:Reference number: PT0222; MUID:91108337; PMID:1899102

C:Accession: PT0267

A:Molecule type: DNA

A:Residues: 1-5 <YAM>

A:Experimental source: B lymphocyte

C:Keywords: heterotetramer; immunoglobulin

Query Match 40.0%; Score 2; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
 ||
 Db 2 GG 3

RESULT 49

D44823

synaptosomal-associated protein SNAP-25 peptide 2 - rabbit (fragment)

N:Alternate names: superprotein peptide 2

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 31-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 15-Jun-1996

C:Accession: D44823

R:Loewy, A.; Liu, W.S.; Baitinger, C.; Willard, M.B.

J. Neurosci. 11, 3412-3421, 1991

A:Title: The major 35S-methionine-labeled rapidly transported protein (superprotein) is

A:Reference number: A44823; MUID:92044785; PMID:1941090

A:Accession: D44823

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-5 <LOE>

A:Experimental source: visual tissue

A>Note: sequence extracted from NCBI backbone (NCBIP:64250)

C:Keywords: membrane trafficking

Query Match 40.0%; Score 2; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
 ||
 Db 3 GS 4

RESULT 50

PT0669

T-cell receptor beta chain V-D-J region (121-3BH) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997

C:Accession: PT0669

R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N region

A:Reference number: PT0509; MUID:91277601; PMID:1711558

C:Accession: PT0669

A>Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-5 <PEE>

A:Experimental source: day 4 postnatal thymus, strain BALB/c

C:Keywords: T-cell receptor

Query Match 40.0%; Score 2; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
 ||
 Db 3 GG 4

RESULT 51

PT0707

T-cell receptor beta chain V-D-J region (140-2H) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997

C:Accession: PT0707; PT0705

R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N region

A:Reference number: PT0509; MUID:91277601; PMID:1711558

C:Accession: PT0707

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-5 <FEE>

A:Experimental source: strain BALB/c (clones 140-2H and 135-1DT)

C:Keywords: T-cell receptor

Query Match 40.0%; Score 2; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
 ||
 Db 3 GG 4

RESULT 52

PT0595

T-cell receptor beta chain V-D-J region (141-1BA) - mouse (fragment)

Species: Mus musculus (house mouse)
 Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 Accession: PT0585
 Feeney, A.J.
 Exp. Med. 174, 115-124, 1991
 Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
 Reference number: PT0509; MUID:91277601; PMID:1711558
 Accession: PT0585
 Status: translation not shown
 Molecule type: mRNA
 Residues: 1-5 <FEE>
 Experimental source: day 19 fetal thymus, strain BALB/c
 Keywords: T-cell receptor

Query Match 40.0%; Score 2; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GG 2
 ||
 4 GG 5

33ULT 53
 r0717
 -cell receptor beta chain V-D-J region - mouse (fragment)
 Species: Mus musculus (house mouse)
 Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 Accession: PT0717; PT0681
 Feeney, A.J.
 Exp. Med. 174, 115-124, 1991
 Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
 Reference number: PT0509; MUID:91277601; PMID:1711558
 Accession: PT0717
 Status: translation not shown
 Molecule type: DNA
 Residues: 1-5 <FEE>
 Experimental source: newborn thymus, strain BALB/c, 140-28
 Accession: PT0681
 Status: translation not shown
 Molecule type: DNA
 Residues: 1-5 <FEE>
 Experimental source: day 18 fetal thymus, strain BALB/c, 154-2C
 Keywords: T-cell receptor

Query Match 40.0%; Score 2; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GG 2
 ||
 4 GG 5

SULT 54
 1512
 riant surface glycoprotein MITa 1.1 - Trypanosoma brucei (fragment)
 Species: Trypanosoma brucei
 Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 07-May-1999
 Accession: B61512
 Holder, A.A.; Cross, G.A.M.
 l. Biochem. Parasitol. 2, 135-150, 1981
 Title: Glycoproteins from variant surface glycoproteins of Trypanosoma brucei. C-termi
 Reference number: A61512; MUID:B1172836; PMID:6163983
 Accession: B61512
 Status: preliminary
 Molecule type: protein
 Residues: 1-6 <HOL>
 Keywords: glycoprotein

Query Match 40.0%; Score 2; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GS 5
 Db 5 GS 6

RESULT 55

A27696
 contraction-inhibiting peptide I - blue mussel
 C;Species: Mytilus edulis (blue mussel)
 C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 08-Dec-1995
 C;Accession: A27696
 R;Hirata, T.; Kubota, I.; Iwasawa, N.; Takabatake, I.; Ikeda, T.; Muneoka, Y.
 Biochem. Biophys. Res. Commun. 152, 1376-1382, 1988
 A;Title: Structures and actions of Mytilus inhibitory peptides.
 A;Reference number: A90142; MUID:88240357; PMID:3377776
 A;Accession: A27696
 A;Molecule type: protein
 A;Residues: 1-6 <HIR>
 C;Keywords: amidated carboxyl end
 F;6/Modified site: amidated carboxyl end (Val) #status experimental

Query Match 40.0%; Score 2; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GS 5
 Db 1 GS 2

RESULT 56

PT0280
 Ig heavy chain CVD3 region (clone 4-91B) - human (fragment)
 C;Species: Homo sapiens (man)
 C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
 R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
 J. Exp. Med. 173, 395-407, 1991
 A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and j.
 A;Reference number: PT0222; MUID:91108337; PMID:1899102
 A;Accession: PT0280
 A;Molecule type: DNA
 A;Residues: 1-6 <YAM>
 A;Experimental source: B lymphocyte
 C;Keywords: heterotetramer; immunoglobulin

Query Match 40.0%; Score 2; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GS 5
 Db 1 GS 2

RESULT 57

PT0514
 T-cell receptor beta chain V-D-J region (100-4AC) - mouse (fragment)
 C;Species: Mus musculus (house mouse)
 C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 C;Accession: PT0514
 R;Feeney, A.J.
 J. Exp. Med. 174, 115-124, 1991
 A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
 A;Reference number: PT0509; MUID:91277601; PMID:1711558
 A;Accession: PT0514
 A;Status: translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-6 <FEE>
 A;Experimental source: adult thymus, strain BALB/c
 C;Keywords: T-cell receptor

```

Query Match      40.0%; Score 2; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
   ||
Db 3 GG 4

RESULT 58
PT0512
T-cell receptor beta chain V-D-J region (100-4D) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0512
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0512
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-6 <FEE>
A:Experimental source: adult thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match      40.0%; Score 2; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
   ||
Db 3 GG 4

RESULT 59
PT0605
T-cell receptor beta chain V-D-J region (120-1L) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0605
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0605
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-6 <FEE>
A:Experimental source: newborn thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match      40.0%; Score 2; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
   ||
Db 5 GG 6

RESULT 60
PT0720
T-cell receptor beta chain V-D-J region (126-1BB) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0548; PT0720
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0548
A:Status: translation not shown

Query Match      40.0%; Score 2; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
   ||
Db 5 GG 6

RESULT 61
PT0560
T-cell receptor beta chain V-D-J region (126-1CE) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0560
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0560
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-6 <FEE>
A:Experimental source: day 18 fetal thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match      40.0%; Score 2; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
   ||
Db 5 GG 6

RESULT 62
PT0723
T-cell receptor beta chain V-D-J region (135-1AF) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0723
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0723
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-6 <FEE>
A:Experimental source: newborn thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match      40.0%; Score 2; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
   ||
Db 5 GG 6

RESULT 63
PT0727
T-cell receptor beta chain V-D-J region (126-1BB) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0548; PT0720
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0548
A:Status: translation not shown

```

T-cell receptor beta chain V-D-J region (161-2F) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 C:Accession: PT0727
 R:Feeney, A.J.
 J. Exp. Med. 174, 115-124, 1991
 A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
 A:Reference number: PT0509; MUID:91277601; PMID:1711558
 A:Accession: PT0727
 A:Status: translation not shown
 A:Residues: 1-6 <PE>
 A:Experimental source: newborn thymus, strain BALB/c
 C:Keywords: T-cell receptor

Query Match 40.0%; Score 2; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 2Y 1 GG 2
 ||
 3 GG 4
 Db

RESULT 64
 PT0730
 T-cell receptor beta chain V-D-J region - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 C:Accession: PT0730; PT0594; PT0731
 R:Feeney, A.J.
 J. Exp. Med. 174, 115-124, 1991
 A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
 A:Reference number: PT0509; MUID:91277601; PMID:1711558
 A:Accession: PT0730
 A:Status: translation not shown
 A:Residues: 1-6 <PE>
 A:Experimental source: newborn thymus, strain BALB/c (clone 163-2C)
 A:Accession: PT0594
 A:Status: translation not shown
 A:Residues: 1-6 <FE2>
 A:Experimental source: newborn thymus, strain BALB/c (clone 165-3E)
 C:Keywords: T-cell receptor

Query Match 40.0%; Score 2; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 2Y 1 GG 2
 ||
 3 GG 4
 Db

RESULT 65
 A41946
 T-cell receptor gamma chain (1t.22) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
 C:Accession: A41946
 R:Whetsell, M.; Mosley, R.L.; Whetsell, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.
 Mol. Cell. Biol. 11, 5902-5909, 1991
 A:Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma gene
 A:Reference number: A41946; MUID:92049316; PMID:1658619
 A:Accession: A41946
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA

Query Match 40.0%; Score 2; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 2Y 1 GG 2
 ||
 5 GG 6
 Db

RESULT 66
 A60224
 Met-enkephalin-Arg-Phe - rabbit
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 21-Oct-1992 #sequence_revision 21-Oct-1992 #text_change 02-Sep-2000
 C:Accession: A60224
 R:Madden IV, J.; Evans, C.J.; Tyler, A.N.; Esch, F.S.; Boehlen, P.; Makk, G.; Weber, E.
 J. Neurochem. 56, 1914-1920, 1991
 A:Title: Isolation and characterization of opioid peptides from rabbit cerebellum.
 A:Reference number: A60224; MUID:91225680; PMID:2027006
 A:Accession: A60224
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-7 <MAD>
 C:Superfamily: proenkephalin
 C:Keywords: neuropeptide; opioid peptide

Query Match 40.0%; Score 2; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 GG 2
 ||
 2 GG 3
 Db

RESULT 67
 S42407
 gramicidin S synthetase component II - Bacillus brevis (fragment)
 C:Species: Bacillus brevis
 C:Date: 20-Oct-1994 #sequence_revision 12-Apr-1996 #text_change 12-Apr-1996
 C:Accession: S42407
 R:Stein, T.; Vater, J.; Kluft, V.; Wittmann-Liebold, B.; Franke, P.; Panico, M.; Mc Dowe
 FEBS Lett. 340, 39-44, 1994
 A:Title: Detection of 4'-phosphopantetheine at the thioester binding site for L-valine o
 A:Reference number: S42407; MUID:94164305; PMID:8119405
 A:Accession: S42407
 A:Molecule type: protein
 A:Residues: 1-7 <STE>

Query Match 40.0%; Score 2; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 GG 2
 ||
 2 GG 3
 Db

RESULT 68
 H33098
 180K exoantigen - malaria parasite (Plasmodium falciparum) (fragments)
 C:Species: Plasmodium falciparum
 C:Date: 24-Aug-1990 #sequence_revision 24-Aug-1990 #text_change 09-Jun-2000
 C:Accession: H33098
 R:Nichols, J.H.; Hager, L.P.
 submitted to the Protein Sequence Database, May 1990
 A:Reference number: A33098
 A:Accession: H33098
 A:Status: preliminary
 A:Molecule type: protein

Query Match 40.0%; Score 2; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 GG 2
 ||
 5 GG 6
 Db

RESULT 69
 A41946
 T-cell receptor gamma chain (1t.22) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
 C:Accession: A41946
 R:Whetsell, M.; Mosley, R.L.; Whetsell, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.
 Mol. Cell. Biol. 11, 5902-5909, 1991
 A:Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma ge
 A:Reference number: A41946; MUID:92049316; PMID:1658619
 A:Accession: A41946
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA

Query Match 40.0%; Score 2; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 GG 2
 ||
 5 GG 6
 Db

RESULT 70
 A41946
 T-cell receptor gamma chain (1t.22) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
 C:Accession: A41946
 R:Whetsell, M.; Mosley, R.L.; Whetsell, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.
 Mol. Cell. Biol. 11, 5902-5909, 1991
 A:Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma ge
 A:Reference number: A41946; MUID:92049316; PMID:1658619
 A:Accession: A41946
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA

A:Residues: 1-6 <WHE>
 C:Keywords: T-cell receptor

Query Match 40.0%; Score 2; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 4 GS 5
 ||
 3 GS 4
 Db

RESULT 66
 A60224
 Met-enkephalin-Arg-Phe - rabbit
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 21-Oct-1992 #sequence_revision 21-Oct-1992 #text_change 02-Sep-2000
 C:Accession: A60224
 R:Madden IV, J.; Evans, C.J.; Tyler, A.N.; Esch, F.S.; Boehlen, P.; Makk, G.; Weber, E.
 J. Neurochem. 56, 1914-1920, 1991
 A:Title: Isolation and characterization of opioid peptides from rabbit cerebellum.
 A:Reference number: A60224; MUID:91225680; PMID:2027006
 A:Accession: A60224
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-7 <MAD>
 C:Superfamily: proenkephalin
 C:Keywords: neuropeptide; opioid peptide

Query Match 40.0%; Score 2; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 GG 2
 ||
 2 GG 3
 Db

RESULT 67
 S42407
 gramicidin S synthetase component II - Bacillus brevis (fragment)
 C:Species: Bacillus brevis
 C:Date: 20-Oct-1994 #sequence_revision 12-Apr-1996 #text_change 12-Apr-1996
 C:Accession: S42407
 R:Stein, T.; Vater, J.; Kluft, V.; Wittmann-Liebold, B.; Franke, P.; Panico, M.; Mc Dowe
 FEBS Lett. 340, 39-44, 1994
 A:Title: Detection of 4'-phosphopantetheine at the thioester binding site for L-valine o
 A:Reference number: S42407; MUID:94164305; PMID:8119405
 A:Accession: S42407
 A:Molecule type: protein
 A:Residues: 1-7 <STE>

Query Match 40.0%; Score 2; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 GG 2
 ||
 2 GG 3
 Db

RESULT 68
 H33098
 180K exoantigen - malaria parasite (Plasmodium falciparum) (fragments)
 C:Species: Plasmodium falciparum
 C:Date: 24-Aug-1990 #sequence_revision 24-Aug-1990 #text_change 09-Jun-2000
 C:Accession: H33098
 R:Nichols, J.H.; Hager, L.P.
 submitted to the Protein Sequence Database, May 1990
 A:Reference number: A33098
 A:Accession: H33098
 A:Status: preliminary
 A:Molecule type: protein

Query Match 40.0%; Score 2; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 GG 2
 ||
 2 GG 3
 Db

RESULT 69
 A41946
 T-cell receptor gamma chain (1t.22) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
 C:Accession: A41946
 R:Whetsell, M.; Mosley, R.L.; Whetsell, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.
 Mol. Cell. Biol. 11, 5902-5909, 1991
 A:Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma ge
 A:Reference number: A41946; MUID:92049316; PMID:1658619
 A:Accession: A41946
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA

Query Match 40.0%; Score 2; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 GG 2
 ||
 5 GG 6
 Db

RESULT 70
 A41946
 T-cell receptor gamma chain (1t.22) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
 C:Accession: A41946
 R:Whetsell, M.; Mosley, R.L.; Whetsell, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.
 Mol. Cell. Biol. 11, 5902-5909, 1991
 A:Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma ge
 A:Reference number: A41946; MUID:92049316; PMID:1658619
 A:Accession: A41946
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA

Query Match 40.0%; Score 2; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 GG 2
 ||
 5 GG 6
 Db

RESULT 71
 A41946
 T-cell receptor gamma chain (1t.22) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
 C:Accession: A41946
 R:Whetsell, M.; Mosley, R.L.; Whetsell, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.
 Mol. Cell. Biol. 11, 5902-5909, 1991
 A:Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma ge
 A:Reference number: A41946; MUID:92049316; PMID:1658619
 A:Accession: A41946
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA

Query Match 40.0%; Score 2; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 GG 2
 ||
 5 GG 6
 Db

RESULT 72
 A41946
 T-cell receptor gamma chain (1t.22) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
 C:Accession: A41946
 R:Whetsell, M.; Mosley, R.L.; Whetsell, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.
 Mol. Cell. Biol. 11, 5902-5909, 1991
 A:Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma ge
 A:Reference number: A41946; MUID:92049316; PMID:1658619
 A:Accession: A41946
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA

Query Match 40.0%; Score 2; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 GG 2
 ||
 5 GG 6
 Db

RESULT 73
 A41946
 T-cell receptor gamma chain (1t.22) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
 C:Accession: A41946
 R:Whetsell, M.; Mosley, R.L.; Whetsell, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.
 Mol. Cell. Biol. 11, 5902-5909, 1991
 A:Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma ge
 A:Reference number: A41946; MUID:92049316; PMID:1658619
 A:Accession: A41946
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA

Query Match 40.0%; Score 2; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 GG 2
 ||
 5 GG 6
 Db

RESULT 74
 A41946
 T-cell receptor gamma chain (1t.22) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
 C:Accession: A41946
 R:Whetsell, M.; Mosley, R.L.; Whetsell, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.
 Mol. Cell. Biol. 11, 5902-5909, 1991
 A:Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma ge
 A:Reference number: A41946; MUID:92049316; PMID:1658619
 A:Accession: A41946
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA

A;Residues: 1-7 <NIC>

Query Match 40.0%; Score 2; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred.No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 4 GS 5
 Db 5 GS 6

RESULT 69
 PT0523
 T-cell receptor beta chain V-D-J region (100-4M) - mouse (fragment)
 C;Species: Mus musculus (house mouse)
 C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 C;Accession: PT0523
 R;Feeney, A.J.
 J. Exp. Med. 174, 115-124, 1991
 A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
 A;Reference number: PT0509; MUID:91277601; PMID:1711558
 A;Accession: PT0523
 A;Status: translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-7 <FEE>
 A;Experimental source: adult thymus, strain BALB/c
 C;Keywords: T-cell receptor

Query Match 40.0%; Score 2; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred.No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GG 2
 Db 5 GG 6

RESULT 70
 PT0642
 T-cell receptor beta chain V-D-J region (111-1H) - mouse (fragment)
 C;Species: Mus musculus (house mouse)
 C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 C;Accession: PT0642
 R;Feeney, A.J.
 J. Exp. Med. 174, 115-124, 1991
 A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
 A;Reference number: PT0509; MUID:91277601; PMID:1711558
 A;Accession: PT0642
 A;Status: translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-7 <FEE>
 A;Experimental source: newborn thymus, strain BALB/c
 C;Keywords: T-cell receptor

Query Match 40.0%; Score 2; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred.No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GG 2
 Db 6 GG 7

RESULT 71
 PT0667
 T-cell receptor beta chain V-D-J region (121-2I) - mouse (fragment)
 C;Species: Mus musculus (house mouse)
 C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 C;Accession: PT0667
 R;Feeney, A.J.
 J. Exp. Med. 174, 115-124, 1991
 A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
 A;Reference number: PT0509; MUID:91277601; PMID:1711558

A;Accession: PT0667
 A;Status: translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-7 <FEE>
 A;Experimental source: day 4 postnatal thymus, strain BALB/c
 C;Keywords: T-cell receptor

Query Match 40.0%; Score 2; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred.No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GG 2
 Db 5 GG 6

RESULT 72
 PT0666
 T-cell receptor beta chain V-D-J region (121-3BD) - mouse (fragment)
 C;Species: Mus musculus (house mouse)
 C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 C;Accession: PT0666
 R;Feeney, A.J.
 J. Exp. Med. 174, 115-124, 1991
 A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions
 A;Reference number: PT0509; MUID:91277601; PMID:1711558
 A;Accession: PT0666
 A;Status: translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-7 <FEE>
 A;Experimental source: day 4 postnatal thymus, strain BALB/c
 C;Keywords: T-cell receptor

Query Match 40.0%; Score 2; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred.No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GG 2
 Db 6 GG 7

RESULT 73
 PT0663
 T-cell receptor beta chain V-D-J region (121-3L) - mouse (fragment)
 C;Species: Mus musculus (house mouse)
 C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 C;Accession: PT0663
 R;Feeney, A.J.
 J. Exp. Med. 174, 115-124, 1991
 A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
 A;Reference number: PT0509; MUID:91277601; PMID:1711558
 A;Accession: PT0663
 A;Status: translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-7 <FEE>
 A;Experimental source: day 4 postnatal thymus, strain BALB/c
 C;Keywords: T-cell receptor

Query Match 40.0%; Score 2; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred.No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GG 2
 Db 6 GG 7

RESULT 74
 PT0542
 T-cell receptor beta chain V-D-J region (126-1BA) - mouse (fragment)
 C;Species: Mus musculus (house mouse)
 C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997

C;Accession: PT0542
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0542

A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-7 <FE>
A;Experimental source: day 18 fetal thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 40.0%; Score 2; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
||
Db 3 GG 4

RESULT 75

PT0683
T-cell receptor beta chain V-D-J region (140-1N) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0683

R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0683

A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-7 <FE>
A;Experimental source: day 18 fetal thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 40.0%; Score 2; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
||
Db 6 GG 7

RESULT 76

PT0719
T-cell receptor beta chain V-D-J region (140-2F) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0719; PT0638

R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0719

A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-7 <FE>
A;Experimental source: newborn thymus, strain BALB/c (clone 140-2F)
A;Accession: PT0638

A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-7 <FE>
A;Experimental source: newborn thymus, strain BALB/c (clone 111-1N)
C;Keywords: T-cell receptor

Query Match 40.0%; Score 2; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
||
Db 6 GG 7

RESULT 77

PT0586
T-cell receptor beta chain V-D-J region (141-1CN) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0586; PT0592
R;Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0586

A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-7 <FE>
A;Experimental source: day 19 fetal thymus, strain BALB/c (clones 141-1CN and 141-1CD)
C;Keywords: T-cell receptor

Query Match 40.0%; Score 2; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
||
Db 6 GG 7

RESULT 78

PT0702
T-cell receptor beta chain V-D-J region (161-2AA) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0702
R;Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0702

A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-7 <FE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 40.0%; Score 2; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
||
Db 6 GG 7

RESULT 79

PT0728
T-cell receptor beta chain V-D-J region (161-2H) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0728
R;Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0728

A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-7 <FE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 40.0%; Score 2; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
 ||
 6 GG 7

Db

RESULT 80
 A38671
 peptidylglycine monooxygenase (EC 1.14.17.3) form 3 - rat (fragment)
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 23-Jun-1993
 C:Accession: A38671
 R:Stoffers, D.A.; Ouafik, L.; Eipper, B.A.
 J. Biol. Chem. 266, 1701-1707, 1991
 A:Title: Characterization of novel mRNAs encoding enzymes involved in peptide alpha-amidation
 A:Reference number: A38671; MUID:91107670; PMID:1988445
 A:Accession: A38671
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-7 <STO>
 C:Keywords: oxidoreductase

Query Match 40.0%; Score 2; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
 ||
 6 GS 7

Db

RESULT 81
 S58797
 serine/threonine-specific protein kinase c-mos - rat (fragment)
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 23-Feb-1997
 C:Accession: S58797
 R:Nagao, Y.
 Biochim. Biophys. Acta 1245, 130-143, 1995
 A:Title: Expression of c-mos protein in cultured rat spermatogenic cells and evidence that it is a protein kinase
 A:Reference number: S58797; MUID:95383384; PMID:7654761
 A:Accession: S58797
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-7 <NAG>
 C:Genetics:
 A:Gene: c-mos
 C:Keywords: phosphotransferase; protein kinase

Query Match 40.0%; Score 2; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2
 ||
 3 GG 4

Db

RESULT 82
 A58718
 carnocin UI49 - Carnobacterium sp. (fragment)
 C:Species: Carnobacterium sp.
 C:Date: 23-Jan-1998 #sequence_revision 30-Jan-1998 #text_change 30-Jan-1998
 C:Accession: A58718
 R:Stoffels, G.; Nissen-Meyer, J.; Gudmundsdottir, A.; Sletten, K.; Holo, H.; Nes, I.F.
 Appl. Environ. Microbiol. 58, 1417-1422, 1992
 A:Title: Purification and characterization of a new bacteriocin isolated from a Carnobacterium
 A:Reference number: A58718; MUID:92321768; PMID:1622206
 A:Accession: A58718

A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-7 <STO>
 C:Keywords: antibiotic; lanthionine

Query Match 40.0%; Score 2; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GS 5
 ||
 1 GS 2

Db

RESULT 83

PC2056

trichodecenen II - fungus (Trichoderma viride)

C:Species: Trichoderma viride

C:Date: 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 20-Oct-2000

C:Accession: PC2056

R:Fujita, T.; Wada, S.; Iida, A.; Nishimura, T.; Kanai, M.; Toyama, N.

Chem. Pharm. Bull. 42, 489-494, 1994

A:Title: Fungal metabolites. XIII. Isolation and structural elucidation of new peptai

A:Reference number: PC2056; MUID:94273211; PMID:8004694

A:Accession: PC2056

A:Status: unencoded polypeptide

A:Molecule type: protein

A:Residues: 1-7 <FUJ>

C:Comment: This protein exhibits activities such as voltage-dependent ion channel for hromaffin cells.

C:Keywords: antibiotic; unencoded polypeptide

F:4/Modified site: 2-amino-2-methylpropanoic acid #status experimental

F:7/Modified site: leucinol #status experimental

Query Match 40.0%; Score 2; DB 4; Length 7;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2

||

1 GG 2

Db

RESULT 84

PC2057

trichodecenen II - fungus (Trichoderma viride)

C:Species: Trichoderma viride

C:Date: 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 20-Oct-2000

C:Accession: PC2057

R:Fujita, T.; Wada, S.; Iida, A.; Nishimura, T.; Kanai, M.; Toyama, N.

Chem. Pharm. Bull. 42, 489-494, 1994

A:Title: Fungal metabolites. XIII. Isolation and structural elucidation of new peptai

A:Reference number: PC2056; MUID:94273211; PMID:8004694

A:Accession: PC2057

A:Status: unencoded polypeptide

A:Molecule type: protein

A:Residues: 1-7 <FUJ>

C:Comment: This protein has unique biological activities such as voltage-dependent io: rom adrenal chromaffin cells.

C:Keywords: antibiotic; unencoded polypeptide

F:4/Modified site: 2-amino-2-methylpropanoic acid #status experimental

F:7/Modified site: leucinol #status experimental

Query Match 40.0%; Score 2; DB 4; Length 7;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG 2

||

1 GG 2

Db

RESULT 85

```

115597
;Species: Homo sapiens (man)
;Date: 04-Jun-1999 #sequence_revision 04-Jun-1999 #text_change 28-Jun-1999
;Accession: S15597
;Title: Characterization of a functional promoter for the human retinoic acid receptor-
;Reference number: S15594; MUID:91088249; PMID:2175878
;Accession: S15597
;Molecule type: DNA
;Residues: 1-7 <BRA>
;Cross-references: EMBL:X55058; NID:935876
;Note: This ORF from Fig. 2 is not annotated in GenBank entry HSRARA2, release 111.0
;Comment: This sequence is not thought to be translated.
;Genetics:
;Gene: GDB:RARA
;Cross-references: GDB:l20337; OMIM:180240
;Map position: 17q12-17q12

Query Match 40.0%; Score 2; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 1 GG 2
b 4 GG 5

RESULT 86
;Species: murine poliovirus (fragment)
;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Feb-1995
;Accession: P10184
;Title: Alteration of amino acid 101 within capsid protein VP-1 changes the pathogenicity
;Reference number: P10184; MUID:90063468; PMID:2479706
;Accession: P10184
;Molecule type: genomic RNA
;Residues: 1-8 <ZUR>
;Keywords: capsid protein

Query Match 40.0%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 1 GG 2
b 2 GG 3

RESULT 87
;Species: Trypanosoma brucei
;Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 20-Mar-1998
;Accession: A21440
;Title: Trypanosome mRNAs share a common 5' spliced leader sequence.
;Reference number: A90853; MUID:84282716; PMID:6088073
;Accession: A21440
;Molecule type: mRNA
;Residues: 1-8 <PAR>
;Cross-references: GB:K02195; NID:gl62150; PID:gl62151
;Keywords: glycoprotein

Query Match 40.0%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 1 GG 2
b 2 GG 3

RESULT 88
;Species: Leucophaea maderae (Madeira cockroach)
;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Jun-2000
;Accession: JS0315
;Title: Isolation, primary structure, and synthesis of leucokinin V and VI: myotropic I
;Reference number: JS0315
;Accession: JS0315
;Molecule type: protein
;Residues: 1-8 <HOL>
;Comment: Leucokinin, a family of cephalomyotropic peptides, stimulate contractile act
;Keywords: amidated carboxyl end; cephalomyotropic peptide
;Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 40.0%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 4 GS 5
b 1 GS 2

RESULT 89
;Species: Pygoscelis adeliae (Adelie penguin)
;Date: 07-Oct-1994 #sequence_revision 04-Nov-1994 #text_change 04-Nov-1994
;Accession: A61467
;Title: Protein Chem. 2, 43-62, 1983
;Reference number: A61467
;Accession: A61467
;Molecule type: protein
;Residues: 1-8 <OSU>
;Comment: Penalbumin is a major protein component of egg whites from penguins but not f
;Keywords: egg white; glycoprotein

Query Match 40.0%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 4 GS 5
b 1 GS 2

RESULT 90
;Species: Homo sapiens (man)
;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
;Accession: P10279
;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and j;
;Reference number: P10222; MUID:91108337; PMID:1899102
;Accession: P10279
;Molecule type: DNA
;Residues: 1-8 <YAM>
;Experimental source: B lymphocyte
;Keywords: heterotetramer; immunoglobulin

Query Match 40.0%; Score 2; DB 2; Length 8;

```

```

Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GG 2
   ||
Db 7 GG 8

RESULT 91
PT0323
Ig heavy chain CRD3 region (clone J2-106B) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C:Accession: PT0323
R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and
A:Reference number: PT0222; MUID:91108337; PMID:1899102
A:Accession: PT0323
A:Molecule type: DNA
A:Residues: 1-8 <YAM>
A:Experimental source: B lymphocyte
C:Keywords: heterotetramer; immunoglobulin

Query Match 40.0%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GS 5
   ||
Db 1 GS 2

RESULT 92
I48934
apolipoprotein A-II - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C:Accession: I48934
R:Ko, M.S.; Wang, X.; Horton, J.H.; Hagen, M.D.; Takahashi, N.; Maezaki, Y.; Nadeau, J.H.
Mamm. Genome 5, 349-355, 1994
A:Title: Genetic mapping of 40 cDNA clones on the mouse genome by PCR.
A:Reference number: I48934; MUID:94319082; PMID:8043949
A:Accession: I48934
A>Status: preliminary; translated from GS/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-8 <RES>
A:Cross-references: EMBL:U05691; NID:g497010; PIDN:AAB60462.1; PID:g642826

Query Match 40.0%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GS 5
   ||
Db 4 GS 5

RESULT 93
A35768
T-cell receptor alpha chain V-J region (34S-281) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 05-Oct-1990 #sequence_revision 05-Oct-1990 #text_change 30-May-1997
C:Accession: A35768
R:Koseki, H.; Imai, K.; Nakayama, F.; Sado, T.; Moriwaki, K.; Taniguchi, M.
Proc. Natl. Acad. Sci. U.S.A. 87, 5248-5252, 1990
A:Title: Homogenous junctional sequence of the VJ4+ T-cell antigen receptor alpha chain
A:Reference number: A35768; MUID:90319089; PMID:2371269
A:Accession: A35768
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-8 <KOS>
A>Note: the sequence shown follows the authors' translation at position 1 of TGT

C:Keywords: T-cell receptor

Query Match 40.0%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GS 5
   ||
Db 7 GS 8

RESULT 94
PT0595
T-cell receptor beta chain V-D-J region (100-2AA) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0595
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0595
A>Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-8 <FEE>
A:Experimental source: newborn thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match 40.0%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GG 2
   ||
Db 7 GG 8

RESULT 95
PT0530
T-cell receptor beta chain V-D-J region (100-4AK) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0530
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0530
A>Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-8 <FEE>
A:Experimental source: adult thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match 40.0%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GG 2
   ||
Db 7 GG 8

RESULT 96
PT0527
T-cell receptor beta chain V-D-J region (100-4I) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0527
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558

```

```

A:Accession: PT0527
A>Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-8 <FEE>
A:Experimental source: adult thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match          40.0%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GG 2
Db 6 GG 7

RESULT 97
PT0639
T-cell receptor beta chain V-D-J region (111-1AA) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0639
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0639
A>Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-8 <FEE>
A:Experimental source: newborn thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match          40.0%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GG 2
Db 5 GG 6

RESULT 98
PT0557
T-cell receptor beta chain V-D-J region (126-1BD) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0557
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0557
A>Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-8 <FEE>
A:Experimental source: day 18 fetal thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match          40.0%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GG 2
Db 6 GG 7

RESULT 99
PT0716
T-cell receptor beta chain V-D-J region (135-1AD) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997

```

```

C:Accession: PT0716; PT0584
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0716
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-8 <FEE>
A:Experimental source: strain BALB/c (clones 135-1AD and 141-1M)
C:Keywords: T-cell receptor

Query Match          40.0%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GG 2
Db 7 GG 8

RESULT 100
PT0725
T-cell receptor beta chain V-D-J region (140-2I) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0725
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0725
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-8 <FEE>
A:Experimental source: newborn thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match          40.0%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GG 2
Db 7 GG 8

Search completed: October 21, 2003, 18:29:09
Job time : 17.1111 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:31:37 ; Search time 29.7222 Seconds
(without alignments)
28.171 Million cell updates/sec

Title: US-10-057-890A-16

Perfect score: 5

Sequence: 1 GGGGS 5

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 629382 seqs, 167460630 residues

Word size : 0

Total number of hits satisfying chosen parameters: 81216

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Listing first 150 summaries

Database : Published Applications_AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*

2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*

3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*

4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*

5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*

6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*

7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*

8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*

9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*

10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*

11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*

12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*

13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*

14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*

15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*

16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*

17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*

18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	5	100.0	5	US-09-287-849-44
2	5	100.0	9	Sequence 44, Appl
3	5	100.0	5	US-09-147-142-31
4	5	100.0	9	Sequence 31, Appl
5	5	100.0	5	US-09-858-616-2
6	5	100.0	9	Sequence 2, Appl
7	5	100.0	5	US-09-778-233-45
8	5	100.0	9	Sequence 45, Appl
9	5	100.0	5	US-09-989-789-3
10	5	100.0	9	Sequence 21, Appl
11	5	100.0	5	US-09-976-787-21
12	5	100.0	9	Sequence 180, App
13	5	100.0	5	US-09-192-854-180
14	5	100.0	9	Sequence 36, Appl
15	5	100.0	5	US-09-761-962-36
16	5	100.0	9	Sequence 5, Appl
17	5	100.0	5	US-09-333-527-5
18	5	100.0	9	Sequence 8, Appl
19	5	100.0	5	US-09-925-796-8
20	5	100.0	9	Sequence 116, App
21	5	100.0	5	US-09-815-837-116
22	5	100.0	9	Sequence 5, Appl
23	5	100.0	5	US-09-033-525-5
24	5	100.0	9	Sequence 7, Appl
25	5	100.0	5	US-09-779-451-7
26	5	100.0	9	Sequence 8, Appl
27	5	100.0	5	US-09-941-450-8

Sequence 25, Appl

Sequence 10, Appl

Sequence 3, Appl

Sequence 22, Appl

Sequence 10, Appl

Sequence 6, Appl

Sequence 123, App

Sequence 3, Appl

Sequence 14, Appl

Sequence 4, Appl

Sequence 56, Appl

Sequence 8, Appl

Sequence 40, Appl

Sequence 1, Appl

Sequence 2, Appl

Sequence 212, App

Sequence 3, Appl

Sequence 8, Appl

Sequence 3, Appl

Sequence 23, Appl

Sequence 84, Appl

Sequence 47, Appl

Sequence 32, Appl

Sequence 81, Appl

Sequence 19, Appl

Sequence 44, Appl

Sequence 42, Appl

Sequence 346, App

Sequence 240, App

Sequence 38, Appl

Sequence 61, Appl

Sequence 15, Appl

Sequence 37, Appl

Sequence 8, Appl

Sequence 22, App

Sequence 207, App

Sequence 1, Appl

Sequence 45, Appl

Sequence 57, Appl

Sequence 45, Appl

Sequence 40, Appl

Sequence 64, Appl

Sequence 3, Appl

Sequence 15, Appl

Sequence 6, Appl

Sequence 18, Appl

Sequence 212, App

Sequence 3, Appl

Sequence 6, Appl

Sequence 16, Appl

Sequence 6, Appl

Sequence 5, Appl

Sequence 4, Appl

Sequence 14, Appl

Sequence 4, Appl

Sequence 23, Appl

Sequence 240, App

Sequence 22, Appl

Sequence 50, Appl

Sequence 1, Appl

Sequence 1, Appl

Sequence 4, Appl

Sequence 47, Appl

Sequence 206, App

Sequence 47, Appl

Sequence 36, Appl

```

89 5 100.0 5 15 US-10-113-424-3
90 5 100.0 5 15 US-10-055-711-6
91 5 100.0 5 15 US-10-270-071-1
92 5 100.0 5 15 US-10-218-131-1
93 5 100.0 5 15 US-10-207-655-39
94 5 100.0 5 15 US-10-288-250-7
95 5 100.0 5 15 US-10-238-102-417
96 5 100.0 6 9 US-09-782-980-134
97 5 100.0 6 10 US-09-883-777-11
98 5 100.0 6 12 US-09-949-039-48
99 5 100.0 6 15 US-10-158-238-7
100 5 100.0 6 15 US-09-815-837-112
101 5 100.0 7 10 US-09-144-886-50
102 5 100.0 7 11 US-09-832-355A-63
103 5 100.0 7 12 US-10-120-508-30
104 5 100.0 7 15 US-10-121-235-16
105 5 100.0 7 15 US-10-112-788-19
106 5 100.0 8 10 US-09-756-283A-16
107 5 100.0 8 14 US-10-081-281-34
108 5 100.0 9 8 US-08-923-852-1
109 5 100.0 9 10 US-09-790-317-5
110 5 100.0 9 12 US-09-930-915A-257
111 5 100.0 9 12 US-10-289-454-344
112 5 100.0 9 12 US-09-833-203-25
113 5 100.0 9 12 US-10-082-014-127
114 5 100.0 9 15 US-10-043-344-144
115 5 100.0 9 15 US-10-043-344-145
116 5 100.0 10 9 US-09-287-849-45
117 5 100.0 10 9 US-09-147-142-30
118 5 100.0 10 9 US-09-848-164-1
119 5 100.0 10 9 US-09-976-787-19
120 5 100.0 10 9 US-09-815-837-128
121 5 100.0 10 10 US-09-865-198-19
122 5 100.0 10 10 US-09-790-317-6
123 5 100.0 10 10 US-09-900-379-1
124 5 100.0 10 11 US-09-572-404B-3627
125 5 100.0 10 11 US-09-572-404B-3706
126 5 100.0 10 11 US-09-572-404B-3707
127 5 100.0 10 11 US-09-572-404B-3708
128 5 100.0 10 11 US-09-572-404B-3709
129 5 100.0 10 11 US-09-572-404B-3789
130 5 100.0 10 11 US-09-572-404B-3821
131 5 100.0 10 11 US-09-572-404B-3823
132 5 100.0 10 11 US-09-572-404B-3825
133 5 100.0 10 11 US-09-572-404B-3827
134 5 100.0 10 11 US-09-572-404B-3829
135 5 100.0 10 11 US-09-572-404B-3831
136 5 100.0 10 11 US-09-572-404B-3831
137 5 100.0 10 11 US-09-572-404B-3963
138 5 100.0 10 11 US-09-572-404B-3964
139 5 100.0 10 11 US-09-572-404B-3965
140 5 100.0 10 11 US-09-572-404B-3983
141 5 100.0 10 11 US-09-572-404B-4158
142 5 100.0 10 11 US-09-572-404B-4159
143 5 100.0 10 11 US-09-572-404B-4189
144 5 100.0 10 11 US-09-798-689-19
145 5 100.0 10 11 US-09-984-010-16
146 5 100.0 10 12 US-09-874-907A-3
147 5 100.0 10 12 US-09-930-915A-256
148 5 100.0 10 12 US-10-262-828A-5
149 5 100.0 10 12 US-09-572-270A-168
150 5 100.0 10 12 US-09-572-270A-426

```

ALIGNMENTS

```

RESULT 1
S-09-287-849-44
Sequence 44, Application US/09287849
Patent No. US2002009459A1
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.

```

```

; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and their Uses
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/09/287,849
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:flexible
; OTHER INFORMATION: polylinker
US-09-287-849-44

Query Match 100.0%; Score 5; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 1 GGGGS 5
|||||

RESULT 2
US-09-147-142-31
; Sequence 31, Application US/09147142
; Patent No. US20020018749A1
; GENERAL INFORMATION:
; APPLICANT: HUDSON, Peter John
; APPLICANT: KORTT, Alex Andrew
; APPLICANT: IRVING, Robert Alexander
; APPLICANT: ATWELL, John Leslie
; TITLE OF INVENTION: HIGH AVIDITY POLYVALENT AND POLYSPECIFIC REAGENTS
; FILE REFERENCE: 016786/0212
; CURRENT APPLICATION NUMBER: US/09/147,142
; CURRENT FILING DATE: 1999-03-05
; EARLIER APPLICATION NUMBER: PCT/AU98/00212
; EARLIER FILING DATE: 1998-03-26
; EARLIER APPLICATION NUMBER: AU PO 5917
; EARLIER FILING DATE: 1997-03-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: peptide linker
US-09-147-142-31

Query Match 100.0%; Score 5; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 1 GGGGS 5
|||||

```


Db 1 GGGGS 5

RESULT 3

US-09-214-645-1
; Sequence 1, Application US/09214645
; Patent No. US20020028443A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; TITLE OF INVENTION: METHOD OF DNA SHUFFLING WITH
; POLYNUCLEOTIDES PRODUCED BY BLOCKING OR INTERRUPTING A
; TITLE OF INVENTION: SYNTHESIS OR AMPLIFICATION PROCESS
; FILE REFERENCE: DIVER1220-2
; CURRENT APPLICATION NUMBER: US/09/214,645
; CURRENT FILING DATE: 1999-09-27
; PRIOR APPLICATION NUMBER: PCT/US97/12239
; PRIOR FILING DATE: 1997-07-09
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: linker peptide
US-09-214-645-1

Query Match 100.0%; Score 5; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|||||

Db 1 GGGGS 5

RESULT 4

US-09-858-616-2
; Sequence 2, Application US/09858616
; Patent No. US20020031771A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: Short, Jay
; TITLE OF INVENTION: SEQUENCE BASED SCREENING
; FILE REFERENCE: DIVER1210-6
; CURRENT APPLICATION NUMBER: US/09/858,616
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US/09/571,499
; PRIOR FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: US/09/557,276
; PRIOR FILING DATE: 2000-04-24
; PRIOR APPLICATION NUMBER: US/08/692,002
; PRIOR FILING DATE: 1996-08-02
; PRIOR APPLICATION NUMBER: US/60/008,317
; PRIOR FILING DATE: 1995-12-07
; PRIOR APPLICATION NUMBER: US/08/944,795
; PRIOR FILING DATE: 1997-10-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Linker peptide
US-09-858-616-2

Query Match 100.0%; Score 5; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|||||

Db 1 GGGGS 5

RESULT 5

US-09-779-233-45
; Sequence 45, Application US/09779233
; Patent No. US20020045158A1
; GENERAL INFORMATION:
; APPLICANT: Case, Casey
; TITLE OF INVENTION: CELLS FOR DRUG DISCOVERY
; FILE REFERENCE: 8325-0010
; CURRENT APPLICATION NUMBER: US/09/779,233
; CURRENT FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: linker
US-09-779-233-45

Query Match 100.0%; Score 5; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|||||

Db 1 GGGGS 5

RESULT 6

US-09-989-789-3
; Sequence 3, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: peptide linker
US-09-989-789-3

Query Match 100.0%; Score 5; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|||||

Db 1 GGGGS 5

RESULT 7

US-09-976-787-21
; Sequence 21, Application US/09976787
; Patent No. US20020064528A1
; GENERAL INFORMATION:
; APPLICANT: Zhu, Zhenping
; APPLICANT: Witte, Larry
; TITLE OF INVENTION: Antibodies Specific to XDR and Uses Thereof
; FILE REFERENCE: 11245/46505
; CURRENT APPLICATION NUMBER: US/09/976,787
; CURRENT FILING DATE: 2001-10-12

PRIOR APPLICATION NUMBER: US 09/493,539
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: US 60/117,726
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 40
SOFTWARE: WordPerfect 8.0 for Windows
SEQ ID NO 21

LENGTH: 5

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: peptide linker

US-09-976-787-21

Query Match 100.0%; Score 5; DB 9; Length 5;
Best Local Similarity 100.0%; Pred.No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGGGS 5

b 1 GGGGS 5

RESULT 8

US-09-192-854-180

Sequence 180, Application US/09192854

Patent No. US20020068276A1

GENERAL INFORMATION:

APPLICANT: Winter, Greg

APPLICANT: Tomlinson, Ian

TITLE OF INVENTION: Methods for Selecting Functional Peptides

FILE REFERENCE: 3789/72916

CURRENT APPLICATION NUMBER: US/09/192,854

CURRENT FILING DATE: 1998-11-17

EARLIER APPLICATION NUMBER: 60/066,729

EARLIER FILING DATE: 1997-11-21

NUMBER OF SEQ ID NOS: 212

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 180

LENGTH: 5

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Linker peptide

OTHER INFORMATION: for connecting variable domains.

US-09-192-854-180

Query Match 100.0%; Score 5; DB 9; Length 5;
Best Local Similarity 100.0%; Pred.No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGGGS 5

b 1 GGGGS 5

RESULT 9

US-09-761-962-36

Sequence 36, Application US/09761962

Patent No. US20020077285A1

GENERAL INFORMATION:

APPLICANT: Memorial Sloan-Kettering Cancer Center

TITLE OF INVENTION: Identification and Characterization of Multiple Splice

TITLE OF INVENTION: Variants of Mu-

TITLE OF INVENTION: opiod Receptor (MOR-1) Gene

FILE REFERENCE: 810002-2000.1

CURRENT APPLICATION NUMBER: US/09/761,962

CURRENT FILING DATE: 2001-01-17

PRIOR APPLICATION NUMBER: 09/743,872

PRIOR FILING DATE: 2001-03-13

NUMBER OF SEQ ID NOS: 46

SOFTWARE: PatentIn version 3.0

SEQ ID NO 36

LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: basic unit of a linking peptide
US-09-761-962-36

Query Match 100.0%; Score 5; DB 9; Length 5;
Best Local Similarity 100.0%; Pred.No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5

Db 1 GGGGS 5

RESULT 10

US-09-333-527-5

Sequence 5, Application US/09333527

Patent No. US20020078472A1

GENERAL INFORMATION:

APPLICANT: Paul CHRISTOU; Eva STROGER; Rainer FISCHER; Carmen MARTIN-VAQUERO; Stef.

TITLE OF INVENTION: METHODS AND MEANS FOR EXPRESSION OF MAMMALIAN POLYPEPTIDES

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fulbright & Jaworski L.L.P.

STREET: 666 Fifth Avenue

CITY: New York City

STATE: New York

COUNTRY: USA

ZIP: 10103

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.25 inch, 1.44mb

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/333,527

FILING DATE: Concurrently Herewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/089,322

FILING DATE: June 15, 1998

ATTORNEY/AGENT INFORMATION:

NAME: Mary Anne Schofield

REGISTRATION NUMBER: 36,669

REFERENCE/DOCKET NUMBER: KL/JIC 202.1 - JEL

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 318-3000

TELEFAX: (212) 752-5958

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 5

TYPE: amino acid

TOPOLOGY: linear

US-09-333-527-5

Query Match 100.0%; Score 5; DB 9; Length 5;
Best Local Similarity 100.0%; Pred.No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5

Db 1 GGGGS 5

RESULT 11

US-09-925-796-8

Sequence 8, Application US/09925796

Patent No. US20020081614A1

GENERAL INFORMATION:

APPLICANT: Case, Casey C.

APPLICANT: Zhang, Lei

```
; APPLICANT: Sangamo Biosciences, Inc.
; TITLE OF INVENTION: Functional Genomics Using Zinc Finger Proteins
; FILE REFERENCE: 019496-002000JUS
; CURRENT APPLICATION NUMBER: US/09/925,796
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/395,448
; PRIOR FILING DATE: 1999-09-14
; PRIOR APPLICATION NUMBER: 09/229,037
; PRIOR FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:linker
US-09-925-796-8

Query Match      100.0%; Score 5; DB 9; Length 5;
Best Local Similarity 100.0%; Pred.No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGGGS 5
      |||||
Db      1 GGGGS 5

RESULT 12
US-09-815-837-116
; Sequence 116, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
; APPLICANT: Carter, Darrick
; APPLICANT: Zhu, Shirley
; APPLICANT: Arimilli, Subhashini
; APPLICANT: Wang, Aijun
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Immune Medators and Related Methods
; FILE REFERENCE: 014058-005670US
; CURRENT APPLICATION NUMBER: US/09/815,837
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 60/191,274
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: US 60/204,249
; PRIOR FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: US 60/264,003
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 116
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:downstream
; OTHER INFORMATION: linker for COS96
US-09-815-837-116
```

```
Query Match      100.0%; Score 5; DB 9; Length 5;
Best Local Similarity 100.0%; Pred.No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGGGS 5
      |||||
Db      1 GGGGS 5
```

```
RESULT 13
US-09-033-525-5
; Sequence 5, Application US/09033525
; Patent No. US2002009037A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Yarkoni, Shai
; APPLICANT: Ben-Yehudah, Ahmi
; APPLICANT: Azar, Yehudith
; APPLICANT: Aqeilan, Rami
; APPLICANT: Belototsky, Ruth
; APPLICANT: Lorberboun-Galski, Haya
; TITLE OF INVENTION: CHIMERIC PROTEINS WITH CELL-TARGETING
; FILE REFERENCE: 9457-009-999
; CURRENT APPLICATION NUMBER: US/09/033,525
; CURRENT FILING DATE: 1998-03-02
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Flexible polylinker
US-09-033-525-5
```

```
Query Match      100.0%; Score 5; DB 9; Length 5;
Best Local Similarity 100.0%; Pred.No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGGGS 5
      |||||
Db      1 GGGGS 5
```

```
RESULT 14
US-09-779-451-7
; Sequence 7, Application US/09779451
; Patent No. US20020094521A1
; GENERAL INFORMATION:
; APPLICANT: Wild, Carl T.
; APPLICANT: Allaway, Graham P.
; TITLE OF INVENTION: Assay for Detection of Viral Fusion Inhibitors
; FILE REFERENCE: 1900.0100003
; CURRENT APPLICATION NUMBER: US/09/779,451
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 60/235,901
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/181,543
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: REPEAT
; LOCATION: (1)-(5)
; OTHER INFORMATION: (GGGGS)x, where x is 1, 2, 3, 4, or 5
; NAME/KEY: misc:feature
; OTHER INFORMATION: Preferred amino acid residues
US-09-779-451-7
```

```
Query Match      100.0%; Score 5; DB 9; Length 5;
Best Local Similarity 100.0%; Pred.No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGGGS 5
      |||||
Db      1 GGGGS 5
```

```
RESULT 15
US-09-941-450-8
; Sequence 8, Application US/09941450
; Patent No. US20020094529A1
; GENERAL INFORMATION:
```

APPLICANT: Case, Casey C.
 TITLE OF INVENTION: GENE IDENTIFICATION
 FILE REFERENCE: S7.US3 / 8325-0007.20
 CURRENT APPLICATION NUMBER: US/09/941,450
 CURRENT FILING DATE: 2001-08-28
 PRIOR APPLICATION NUMBER: 09/395,448
 PRIOR FILING DATE: 1999-09-14
 NUMBER OF SEQ ID NOS: 23
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 8
 LENGTH: 5
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence:linker

IS-09-941-450-8

Query Match 100.0%; Score 5; DB 9; Length 5;
 Best Local Similarity 100.0%; Pred. No. 5.7e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 1 GGGGS 5
 |||||
 b 1 GGGGS 5

RESULT 16

IS-09-818-247-25
 Sequence 25, Application US/09818247
 Patent No. US20020102657A1
 GENERAL INFORMATION:
 APPLICANT: Mostov, Keith E.
 APPLICANT: Chapin, Steven J.
 APPLICANT: Richman-Bisenat, Janice
 TITLE OF INVENTION: The Regents of the University of California
 TITLE OF INVENTION: Ligands Directed to the No. US20020102657A1-Secretory Component,
 FILE REFERENCE: 18062E-00091005
 CURRENT APPLICATION NUMBER: US/09/818,247
 CURRENT FILING DATE: 2001-03-26
 PRIOR APPLICATION NUMBER: WO PCT/US01/09699
 PRIOR FILING DATE: 2001-03-26
 PRIOR APPLICATION NUMBER: US 60/192,197
 PRIOR FILING DATE: 2000-03-27
 PRIOR APPLICATION NUMBER: US 60/192,198
 PRIOR FILING DATE: 2000-03-27
 NUMBER OF SEQ ID NOS: 26
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 25
 LENGTH: 5
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence:peptide linker

S-09-818-247-25

Query Match 100.0%; Score 5; DB 10; Length 5;
 Best Local Similarity 100.0%; Pred. No. 5.7e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 1 GGGGS 5
 |||||
 b 1 GGGGS 5

ESULT 17

S-09-883-777-10
 Sequence 10, Application US/09883777
 Patent No. US20020110853A1
 GENERAL INFORMATION:
 APPLICANT: Wiley, Steven R.
 TITLE OF INVENTION: TWEAK RECEPTOR

FILE REFERENCE: 2968-C
 CURRENT APPLICATION NUMBER: US/09/883,777
 CURRENT FILING DATE: 2001-06-18
 PRIOR APPLICATION NUMBER: US 60/172,878
 PRIOR FILING DATE: 1999-12-20
 PRIOR APPLICATION NUMBER: US 60/203,347
 PRIOR FILING DATE: 2000-05-10
 PRIOR APPLICATION NUMBER: PCT/US00/34755
 PRIOR FILING DATE: 2000-12-19
 PRIOR APPLICATION NUMBER: US 09/742,454
 PRIOR FILING DATE: 2000-12-19
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 10
 LENGTH: 5
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Linker Moiety
 US-09-883-777-10

Query Match 100.0%; Score 5; DB 10; Length 5;
 Best Local Similarity 100.0%; Pred. No. 5.7e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
 |||||
 Db 1 GGGGS 5

RESULT 18

US-09-867-262-3
 Sequence 3, Application US/09867262
 Patent No. US20020119457A1
 GENERAL INFORMATION:
 APPLICANT: DIVERSA CORPORATION
 APPLICANT: SHORT, Jay
 APPLICANT: PREY, Gerhard
 TITLE OF INVENTION: END SELECTION IN DIRECTED EVOLUTION
 FILE REFERENCE: DEVER1460-17
 CURRENT APPLICATION NUMBER: US/09/867,262
 CURRENT FILING DATE: 2001-05-29
 PRIOR APPLICATION NUMBER: US 09/267,118
 PRIOR FILING DATE: 1999-03-09
 PRIOR APPLICATION NUMBER: US 09/246,178
 PRIOR FILING DATE: 1999-02-04
 PRIOR APPLICATION NUMBER: US 09/185,373
 PRIOR FILING DATE: 1998-11-03
 PRIOR APPLICATION NUMBER: US 08/760,489
 PRIOR FILING DATE: 1996-12-05
 PRIOR APPLICATION NUMBER: US 60/008,311
 PRIOR FILING DATE: 1995-12-07
 PRIOR APPLICATION NUMBER: US 08/962,504
 PRIOR FILING DATE: 1997-10-31
 PRIOR APPLICATION NUMBER: US 08/677,112
 PRIOR APPLICATION NUMBER: US 08/651,568
 PRIOR FILING DATE: 1996-05-22
 PRIOR APPLICATION NUMBER: US 60/008,316
 PRIOR FILING DATE: 1995-12-07
 NUMBER OF SEQ ID NOS: 14
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 3
 LENGTH: 5
 TYPE: PRT
 ORGANISM: Artificial sequence
 FEATURE:
 OTHER INFORMATION: Antibody spacer peptide. The entire peptide sequence can be
 OTHER INFORMATION: repeated more than one time
 US-09-867-262-3

Query Match 100.0%; Score 5; DB 10; Length 5;
 Best Local Similarity 100.0%; Pred. No. 5.7e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|
|
|
|
|
Db 1 GGGGS 5

RESULT 19

US-09-780-933-22
; Sequence 22, Application US/09780933
; Patent No. US20020127652A1
; GENERAL INFORMATION:
; APPLICANT: SCHAMBEY, HANS T.
; APPLICANT: ANDERSEN, KIM V.
; APPLICANT: VAN DEN HAZEL, BART
; APPLICANT: CHRISTIANSEN, JESPER
; APPLICANT: JEPPESEN, CLAUS B.
; TITLE OF INVENTION: FOLLICLE STIMULATING HORMONES
; FILE REFERENCE: 31-001000US
; CURRENT APPLICATION NUMBER: US/09/780,933
; CURRENT FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: PA 2000 00220
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 60/184,035
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: PA 2000 01092
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,558
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-780-933-22

Query Match 100.0%; Score 5; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|
|
|
|
|
Db 1 GGGGS 5

RESULT 20

US-09-480-236-10
; Sequence 10, Application US/09480236
; Patent No. US20020142000A1
; GENERAL INFORMATION:
; APPLICANT: Digan, Mary Ellen
; APPLICANT: Lake, Philip
; APPLICANT: Wright, Richard M.
; TITLE OF INVENTION: Anti-CD3 Immunotoxins and Therapeutic Uses Therefor
; FILE REFERENCE: CGC 4-31157A/USN
; CURRENT APPLICATION NUMBER: US/09/480,236
; CURRENT FILING DATE: 2000-01-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: peptide linker
US-09-480-236-10

Query Match 100.0%; Score 5; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|
|
|
|
|
Db 1 GGGGS 5

RESULT 21

US-09-731-558-6
; Sequence 6, Application US/09731558
; Patent No. US20020146691A1
; GENERAL INFORMATION:
; APPLICANT: Case, Casey Christopher
; APPLICANT: Liu, Qiang
; APPLICANT: Rebar, Edward J.
; APPLICANT: Sangamo Biosciences, Inc.
; TITLE OF INVENTION: Methods of Using Randomized Libraries of Zinc Finger
; FILE REFERENCE: 019496-003210US
; CURRENT APPLICATION NUMBER: US/09/731,558
; CURRENT FILING DATE: 2000-12-06
; PRIOR APPLICATION NUMBER: US 09/456,100
; PRIOR FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: polypeptide
; OTHER INFORMATION: linker
US-09-731-558-6

Query Match 100.0%; Score 5; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|
|
|
|
|
Db 1 GGGGS 5

RESULT 22

US-09-828-708-123
; Sequence 123, Application US/09828708
; Patent No. US20020146753A1
; GENERAL INFORMATION:
; APPLICANT: Ditzel, H.
; APPLICANT: Burton, D.
; APPLICANT: Schaller, M.
; TITLE OF INVENTION: Autoantibodies to glucose-6-phosphate isomerase and their part
; FILE REFERENCE: 1361.005US1
; CURRENT APPLICATION NUMBER: US/09/828,708
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 123
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: A flexible five amino acid tether.
US-09-828-708-123

Query Match 100.0%; Score 5; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|
|
|
|
|
Db 1 GGGGS 5

```

US-09-885-551A-3
; Sequence 3, Application US/09885551A
; Patent No. US20020146762A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: DJAVAKHISHVILI, Tsotne
; APPLICANT: FREY, Gerhard
; TITLE OF INVENTION: EXONUCLEASE-MEDIATED NUCLEIC ACID REASSEMBLY IN
; TITLE OF INVENTION: DIRECTED EVOLUTION
; FILE REFERENCE: DIVER1460-14
; CURRENT APPLICATION NUMBER: US/09/885,551A
; CURRENT FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US/09/535,754
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 09/522,289
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Antibody spacer peptide. The entire peptide sequence
; OTHER INFORMATION: can be repeated more than one time
; OTHER INFORMATION: ated more than one time
US-09-885-551A-3

Query Match 100.0%; Score 5; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y 1 GGGGS 5
|||||
2B 1 GGGGS 5

RESULT 24
JS-09-756-283A-14
; Sequence 14, Application US/09756283A
; Patent No. US20020151478A1
; GENERAL INFORMATION:
; APPLICANT: Chernajovsky, Yuri
; APPLICANT: Dreja, Hanna Stina
; APPLICANT: Adams, Gillian
; TITLE OF INVENTION: Latent Fusion Protein
; FILE REFERENCE: 0623.1000000
; CURRENT APPLICATION NUMBER: US/09/756,283A
; CURRENT FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Flexible linker
JS-09-756-283A-14

Query Match 100.0%; Score 5; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y 1 GGGGS 5
|||||
2B 1 GGGGS 5

RESULT 25
US-09-144-886-4
; Sequence 4, Application US/09144886
; Patent No. US20020155114A1
; GENERAL INFORMATION:
; APPLICANT: Marks, James D
; APPLICANT: Amersdorfer, Peter
; TITLE OF INVENTION: Therapeutic Monoclonal Antibodies That Neutralize
; TITLE OF INVENTION: Botulinum Neurotoxins
; FILE REFERENCE: 2500.117USO
; CURRENT APPLICATION NUMBER: US/09/144,886
; CURRENT FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Gly4Ser linker
US-09-144-886-4

Query Match 100.0%; Score 5; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|||||
DB 1 GGGGS 5

RESULT 26
US-09-999-745-56
; Sequence 56, Application US/09999745
; Patent No. US20020157120A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: CIRCULARLY PERMUTED FLUORESCENT PROTEIN INDICATORS
; FILE REFERENCE: REGEN1470-1
; CURRENT APPLICATION NUMBER: US/09/999,745
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 09/316,920
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 56
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Linker moiety; sequence repeated indefinitely
US-09-999-745-56

Query Match 100.0%; Score 5; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|||||
DB 1 GGGGS 5

RESULT 27
US-09-942-087A-8
; Sequence 8, Application US/09942087A
; Patent No. US20020160940A1
; GENERAL INFORMATION:
; APPLICANT: Case, Casey Christopher
; APPLICANT: Wolffe, Alan
; APPLICANT: Iurnov, Fyodor
; APPLICANT: Lai, Albert
; APPLICANT: Snowden, Andrew

```

```

; APPLICANT: Tan, Siyuan
; APPLICANT: Gregory, Philip
; TITLE OF INVENTION: MODULATION OF ENDOGENOUS GENE EXPRESSION IN CELLS
; FILE REFERENCE: 8325-0002.21 / S2-US5
; CURRENT APPLICATION NUMBER: US/09/942,087A
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 09/229,037
; PRIOR FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:linker
US-09-942-087A-8

Query Match      100.0%; Score 5; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGGGS 5
Db      1 GGGGS 5

RESULT 28
US-09-942-090-8
; Sequence 8, Application US/09942090
; Patent No. US20020164575A1
; GENERAL INFORMATION:
; APPLICANT: Case, Casey C.
; APPLICANT: Zhang, Lei
; APPLICANT: Sangamo Biosciences, Inc.
; TITLE OF INVENTION: Functional Genomics Using Zinc Finger Proteins
; FILE REFERENCE: 019496-0020000US
; CURRENT APPLICATION NUMBER: US/09/942,090
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 09/395,448
; PRIOR FILING DATE: 1999-09-14
; PRIOR APPLICATION NUMBER: 09/229,007
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 09/229,037
; PRIOR FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:linker
US-09-942-090-8

Query Match      100.0%; Score 5; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGGGS 5
Db      1 GGGGS 5

RESULT 29
US-09-554-000-40
; Sequence 40, Application US/09554000
; Patent No. US20020165364A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; FILE REFERENCE: DETECTION OF ANALYTES

```

```

; FILE REFERENCE: 07257/042001
; CURRENT APPLICATION NUMBER: US/09/554,000
; CURRENT FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 08/818,252
; PRIOR FILING DATE: 1997-03-14
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Linker moiety: sequence repeated indefinitely
US-09-554-000-40

Query Match      100.0%; Score 5; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGGGS 5
Db      1 GGGGS 5

RESULT 30
US-09-792-793A-1
; Sequence 1, Application US/09792793A
; Patent No. US20020168370A1
; GENERAL INFORMATION:
; APPLICANT: McDonald, John R.
; APPLICANT: Coggins, Philip
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE
; TITLE OF INVENTION: OTHER INFLAMMATORY CONDITIONS AND DISORDERS
; FILE REFERENCE: 25020-601D
; CURRENT APPLICATION NUMBER: US/09/792,793A
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 5
; TYPE: PRT
; ORGANISM: homo sapien
; FEATURE:
; OTHER INFORMATION: peptide linker for use in conjugates
US-09-792-793A-1

Query Match      100.0%; Score 5; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGGGS 5
Db      1 GGGGS 5

RESULT 31
US-09-792-793A-2
; Sequence 2, Application US/09792793A
; Patent No. US20020168370A1
; GENERAL INFORMATION:
; APPLICANT: McDonald, John R.
; APPLICANT: Coggins, Philip
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE
; TITLE OF INVENTION: OTHER INFLAMMATORY CONDITIONS AND DISORDERS
; FILE REFERENCE: 25020-601D
; CURRENT APPLICATION NUMBER: US/09/792,793A
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 5
; TYPE: PRT
; ORGANISM: homo sapien

```

```
FEATURE:
NAME/KEY: REPEAT
LOCATION: (1)...(5)
OTHER INFORMATION: peptide linker for use in conjugates
JS-09-792-793A-2
Query Match 100.0%; Score 5; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 1 GGGGS 5

RESULT 32
JS-09-846-033B-212
Sequence 212, Application US/09846033B
Publication No. US2003004404A1
GENERAL INFORMATION:
APPLICANT: Rebar, Edward
APPLICANT: Jamieson, Andrew
APPLICANT: Liu, Qiang
APPLICANT: Liu, Pei-Qi
APPLICANT: Wolffe, Alan
APPLICANT: Eisenberg, Stephen P.
APPLICANT: Jarvis, Eric
APPLICANT: Sangamo BioSciences, Inc.
TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
TITLE OF INVENTION: Finger Proteins
FILE REFERENCE: 019496-005820US
CURRENT APPLICATION NUMBER: US/09/846.033B
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: US 09/733,604
PRIOR FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 09/736,083
PRIOR FILING DATE: 2000-12-12
NUMBER OF SEQ ID NOS: 252
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 212
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: peptide linker
JS-09-846-033B-212
Query Match 100.0%; Score 5; DB 11; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 1 GGGGS 5

RESULT 33
JS-09-990-186-3
Sequence 3, Application US/09990186
Publication No. US20030068675A1
GENERAL INFORMATION:
APPLICANT: Liu, Qiang
TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
TITLE OF INVENTION: TRIPLETS BY ZINC FINGERS
FILE REFERENCE: 8325-0011.21 / S11-US3
CURRENT APPLICATION NUMBER: US/09/990.186
CURRENT FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 4085
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: peptide linker
JS-09-989-994-3
Query Match 100.0%; Score 5; DB 11; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 1 GGGGS 5

RESULT 34
US-09-897-844-8
Sequence 8, Application US/09897844
Publication No. US20030087817A1
GENERAL INFORMATION:
APPLICANT: Cox III, George No. US20030087817Albert
APPLICANT: Case, Casey Christopher
APPLICANT: Eisenberg, Stephen P.
APPLICANT: Jarvis, Eric Edward
APPLICANT: Spratt, Sharon Kaye
APPLICANT: Sangamo BioSciences, Inc.
TITLE OF INVENTION: Regulation of Endogenous Gene Expression in Cells Using
TITLE OF INVENTION: Zinc Finger Proteins
FILE REFERENCE: 019496-002200US
CURRENT APPLICATION NUMBER: US/09/897.844
CURRENT FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: 09/229,037
PRIOR FILING DATE: 1999-01-12
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: linker
US-09-897-844-8
Query Match 100.0%; Score 5; DB 11; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 1 GGGGS 5

RESULT 35
US-09-989-994-3
Sequence 3, Application US/09899994
Publication No. US20030104526A1
GENERAL INFORMATION:
APPLICANT: Liu, Qiang
TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
TITLE OF INVENTION: TRIPLETS BY ZINC FINGERS
FILE REFERENCE: 8325-0011.20 / S11-US2
CURRENT APPLICATION NUMBER: US/09/989.994
CURRENT FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 4085
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: peptide linker
US-09-989-994-3
Query Match 100.0%; Score 5; DB 11; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
```


Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
 Db 1 GGGGS 5

RESULT 36
 US-09-911-261A-23
 ; Sequence 23, Application US/09911261A
 ; Publication No. US20030134350A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sera, Takashi
 ; TITLE OF INVENTION: Zinc Finger Domain Recognition Code and Uses Thereof
 ; FILE REFERENCE: 109845.135
 ; CURRENT APPLICATION NUMBER: US/09/911,261A
 ; CURRENT FILING DATE: 2001-07-23
 ; PRIOR APPLICATION NUMBER: US 60/220,060
 ; PRIOR FILING DATE: 2000-07-21
 ; NUMBER OF SEQ ID NOS: 69
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 23
 ; LENGTH: 5
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Flexible linker
 US-09-911-261A-23

Query Match 100.0%; Score 5; DB 12; Length 5;
 Best Local Similarity 100.0%; Pred. No. 5.7e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
 Db 1 GGGGS 5

RESULT 37
 US-09-942-024-84
 ; Sequence 84, Application US/09942024
 ; Publication No. US20030143650A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Steward, Lance E.
 ; APPLICANT: Fernandez-Salas, Ester
 ; APPLICANT: Aoki, Kei Roger
 ; TITLE OF INVENTION: Fret Protease Assays For Botulinum
 ; FILE REFERENCE: P-AR 4803
 ; CURRENT APPLICATION NUMBER: US/09/942,024
 ; CURRENT FILING DATE: 2001-08-28
 ; NUMBER OF SEQ ID NOS: 96
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 84
 ; LENGTH: 5
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: synthetic construct
 US-09-942-024-84

Query Match 100.0%; Score 5; DB 12; Length 5;
 Best Local Similarity 100.0%; Pred. No. 5.7e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
 Db 1 GGGGS 5

RESULT 38
 US-09-942-098-84
 ; Sequence 84, Application US/09942098

; Publication No. US20030143651A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Steward, Lance E.
 ; APPLICANT: Fernandez-Salas, Ester
 ; APPLICANT: Aoki, Kei Roger
 ; TITLE OF INVENTION: Fret Protease Assays For Clostridial
 ; TITLE OF INVENTION: Toxins
 ; FILE REFERENCE: P-AR 4802
 ; CURRENT APPLICATION NUMBER: US/09/942,098
 ; CURRENT FILING DATE: 2001-08-28
 ; NUMBER OF SEQ ID NOS: 96
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 84
 ; LENGTH: 5
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: synthetic construct
 US-09-942-098-84

Query Match 100.0%; Score 5; DB 12; Length 5;
 Best Local Similarity 100.0%; Pred. No. 5.7e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
 Db 1 GGGGS 5

RESULT 39
 US-10-244-821-47
 ; Sequence 47, Application US/10244821
 ; Publication No. US2003014323A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Goshorn, Stephen Charles
 ; APPLICANT: Graves, Scott Stoll
 ; APPLICANT: Schultz, Joanne Elaine
 ; APPLICANT: Lin, Yukang
 ; APPLICANT: Sanderson, James Allen
 ; APPLICANT: Reno, John M.
 ; APPLICANT: Dearstyne, Erica A.
 ; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
 ; TITLE OF INVENTION: METHODS OF USE THEREOF
 ; FILE REFERENCE: 690022.547C3
 ; CURRENT APPLICATION NUMBER: US/10/244,821
 ; CURRENT FILING DATE: 2002-09-16
 ; NUMBER OF SEQ ID NOS: 92
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 47
 ; LENGTH: 5
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Gly Ser linker
 US-10-244-821-47

Query Match 100.0%; Score 5; DB 12; Length 5;
 Best Local Similarity 100.0%; Pred. No. 5.7e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
 Db 1 GGGGS 5

RESULT 40
 US-10-262-630-32
 ; Sequence 32, Application US/10262630
 ; Publication No. US20030138401A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dahiyat, Bassil I.
 ; APPLICANT: Desjarlais, John R.
 ; APPLICANT: Filikov, Anton

APPLICANT: Muchhal, Umesh
APPLICANT: Tansey, Malu Lourdas G.
APPLICANT: Zalevsky, Jonathan
TITLE OF INVENTION: PROTEIN BASED TNF-ALPHA VARIANTS FOR THE TREATMENT OF TNF-ALPHA
FILE REFERENCE: A-68990-4/RFT/RMS/RMK
CURRENT APPLICATION NUMBER: US/10/262,630
CURRENT FILING DATE: 2003-01-27
PRIOR APPLICATION NUMBER: US 60/186,427
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: US 09/945,150
PRIOR FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: US 09/798,789
PRIOR FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: US 09/981,289
PRIOR FILING DATE: 2001-10-15
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.2
SEQ ID NO 32
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: linker sequence
JS-10-262-630-32

Query Match 100.0%; Score 5; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y 1 GGGGS 5
|||||
2B 1 GGGGS 5

RESULT 41
JS-10-214-473-81
Sequence 81, Application US/10214473
Publication No. US2003014843A1
GENERAL INFORMATION:
APPLICANT: Zhao, Lishan
APPLICANT: Mathur, Eric J.
APPLICANT: Weiner, David
APPLICANT: Richardson, Toby
APPLICANT: Milan, Aileen
APPLICANT: Burk, Mark J.
APPLICANT: Han, Bin
APPLICANT: Short, Jay M.
TITLE OF INVENTION: EPOXIDE HYDROLASES, NUCLEIC ACIDS ENCODING
FILE REFERENCE: 09010-600001
CURRENT APPLICATION NUMBER: US/10/214,473
CURRENT FILING DATE: 2002-08-05
PRIOR APPLICATION NUMBER: US 60/309,478
PRIOR FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: US 60/393,378
PRIOR FILING DATE: 2002-07-03
NUMBER OF SEQ ID NOS: 83
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 81
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: linker peptide
JS-10-214-473-81

Query Match 100.0%; Score 5; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y 1 GGGGS 5
|||||

Db 1 GGGGS 5

RESULT 42
US-10-270-073-19
Sequence 19, Application US/10270073
Publication No. US20030148409A1
GENERAL INFORMATION:
APPLICANT: ROSSI, EDMUND
APPLICANT: CHANG, CHIEN-HSING KEN
APPLICANT: GOLDENBERG, DAVID
TITLE OF INVENTION: DIRECT TARGETING BINDING PROTEINS
FILE REFERENCE: 042418/0112
CURRENT APPLICATION NUMBER: US/10/270,073
CURRENT FILING DATE: 2002-10-15
PRIOR APPLICATION NUMBER: 60/328,835
PRIOR FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 60/341,881
PRIOR FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 60/345,641
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: 60/404,919
PRIOR FILING DATE: 2002-08-22
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 19
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic Gly-Ser linker
US-10-270-073-19

Query Match 100.0%; Score 5; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|||||
Db 1 GGGGS 5

RESULT 43
US-10-359-460-44
Sequence 44, Application US/10359460
Publication No. US20030147911A1
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, Davin C.
APPLICANT: Alderson, Mark
APPLICANT: Campos-Neto, Antonio
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
FILE REFERENCE: 014058-009020US
CURRENT APPLICATION NUMBER: US/10/359,460
CURRENT FILING DATE: 2003-02-05
PRIOR APPLICATION NUMBER: US/09/287,849
PRIOR FILING DATE: 1999-04-07
PRIOR APPLICATION NUMBER: US 08/818,112
PRIOR FILING DATE: 1997-03-13
PRIOR APPLICATION NUMBER: US 08/942,578
PRIOR FILING DATE: 1997-10-01
PRIOR APPLICATION NUMBER: US 09/025,197
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 09/056,556
PRIOR FILING DATE: 1998-04-07
PRIOR APPLICATION NUMBER: US 09/223,040
PRIOR FILING DATE: 1998-12-30
NUMBER OF SEQ ID NOS: 46
SOFTWARE: PatentIn Ver. 2.1

```
; SEQ ID NO 44
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:flexible
US-10-359-460-44

Query Match      100.0%; Score 5; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGGGS 5
Db      1 GGGGS 5
      |||||

RESULT 44
US-10-072-301-42
; Sequence 42, Application US/10072301
; Publication No. US20030152913A1
; GENERAL INFORMATION:
; APPLICANT: Hua, Shao-bing
; APPLICANT: Pauling, Michelle H.
; APPLICANT: Zhu, Li
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODY AGAINST CORECEPTORS FOR HUMAN IMMUNODEFICIENCY VIRUS
; FILE REFERENCE: 25636-718
; CURRENT APPLICATION NUMBER: US/10/072,301
; CURRENT FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Linker Sequence
US-10-072-301-42

Query Match      100.0%; Score 5; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGGGS 5
Db      1 GGGGS 5
      |||||

RESULT 45
US-10-289-454-346
; Sequence 346, Application US/10289454
; Publication No. US20030157479A1
; GENERAL INFORMATION:
; APPLICANT: Bachmann, Martin
; APPLICANT: Jennings, Gary
; APPLICANT: Sonderegger, Ivo
; TITLE OF INVENTION: Antigen Arrays for Treatments of Allergic Eosinophilic Diseases
; FILE REFERENCE: 1700.0360001
; CURRENT APPLICATION NUMBER: US/10/289,454
; CURRENT FILING DATE: 2003-02-10
; PRIOR APPLICATION NUMBER: US 60/396,636
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: PCT/IB02/00166
; PRIOR FILING DATE: 2002-01-21
; PRIOR APPLICATION NUMBER: US 10/050,902
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/331,045
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 386
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 346
```

```
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Glycine serine linkers
US-10-289-454-346

Query Match      100.0%; Score 5; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGGGS 5
Db      1 GGGGS 5
      |||||

RESULT 46
US-10-289-660-240
; Sequence 240, Application US/10289660
; Publication No. US20030157561A1
; GENERAL INFORMATION:
; APPLICANT: KOLKMAN, JOOST A.
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: GOVINDARAJAN, SRIDHAR
; TITLE OF INVENTION: COMBINATORIAL LIBRARIES OF MONOMER DOMAINS
; FILE REFERENCE: 0319.510US
; CURRENT APPLICATION NUMBER: US/10/289,660
; CURRENT FILING DATE: 2003-11-06
; PRIOR APPLICATION NUMBER: 10/133,128
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: 60/374,107
; PRIOR FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: 60/333,359
; PRIOR FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 60/337,209
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/286,823
; PRIOR FILING DATE: 2001-04-26
; NUMBER OF SEQ ID NOS: 244
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 240
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide linker
US-10-289-660-240

Query Match      100.0%; Score 5; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGGGS 5
Db      1 GGGGS 5
      |||||

RESULT 47
US-09-969-748C-38
; Sequence 38, Application US/09969748C
; Publication No. US20030161809A1
; GENERAL INFORMATION:
; APPLICANT: ARIZEXE PHARMACEUTICALS, INC.
; APPLICANT: HOUSTON, Lou, L.
; APPLICANT: SHERIDAN, Philip, J.
; APPLICANT: HAWLEY, Stephen
; APPLICANT: GLYNN, Jacqueline, M.
; APPLICANT: CHAPIN, Steven
; APPLICANT: BASU, Amaresh
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TRANSPORT OF BIOLOGICALLY ACTIVE AGENTS ACROSS CELLULAR BARRIERS
; FILE REFERENCE: 057220-0303
; CURRENT APPLICATION NUMBER: US/09/969,748C
; CURRENT FILING DATE: 2002-12-10
```

PRIOR APPLICATION NUMBER: US 60/267,601
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: US 60/248,819
PRIOR FILING DATE: 2000-11-14
PRIOR APPLICATION NUMBER: US 60/248,478
PRIOR FILING DATE: 2000-11-13
PRIOR APPLICATION NUMBER: US 60/237,929
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 115
SOFTWARE: PatentIn version 3.0
SEQ ID NO 38
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: linker sequence
S-09-969-748C-38

Query Match 100.0%; Score 5; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGGGS 5
|
|
|
|
b 1 GGGGS 5

RESULT 48

US-09-992-124A-61
Sequence 61, Application US/09992124A
Publication No. US20030162289A1
GENERAL INFORMATION:
APPLICANT: Heidaran, Mohammad A.
APPLICANT: Haaland, Perry D.
APPLICANT: Wilkins, Jamie H.
APPLICANT: Spargo, Catherine A.
APPLICANT: Campbell, Robert L.
TITLE OF INVENTION: Peptides Promoting Cell Adherence, Growth and Secretion
FILE REFERENCE: 102-410
CURRENT APPLICATION NUMBER: US/09/992,124A
CURRENT FILING DATE: 2002-05-21
NUMBER OF SEQ ID NOS: 70
SOFTWARE: PatentIn version 3.1
SEQ ID NO 61
LENGTH: 5
TYPE: PRT
ORGANISM: artificial sequence
FEATURE:
OTHER INFORMATION: Synthetic peptide selected for biological activity
S-09-992-124A-61

Query Match 100.0%; Score 5; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGGGS 5
|
|
|
|
b 1 GGGGS 5

RESULT 49

S-10-328-190-15
Sequence 15, Application US/10328190
Publication No. US20030162709A1
GENERAL INFORMATION:
APPLICANT: ROSSI, EDMUND A.
APPLICANT: CHANG, CHIEN-HSING KEN
APPLICANT: GOLDENBERG, DAVID M.
TITLE OF INVENTION: METHODS OF GENERATING MULTISPECIFIC, MULTIVALENT AGENTS
TITLE OF INVENTION: FROM VH AND VL DOMAINS
FILE REFERENCE: 042418/0115
CURRENT APPLICATION NUMBER: US/10/328,190
CURRENT FILING DATE: 2003-04-03

PRIOR APPLICATION NUMBER: 60/342,103
PRIOR FILING DATE: 2001-12-26
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 15
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: linker peptide
US-10-328-190-15

Query Match 100.0%; Score 5; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|
|
|
|
Db 1 GGGGS 5

RESULT 50

US-09-949-039-37
Sequence 37, Application US/09949039
Publication No. US20030166160A1
GENERAL INFORMATION:
APPLICANT: HAWLEY, STEPHEN B.
TITLE OF INVENTION: COMPOUNDS AND MOLECULAR COMPLEXES COMPRISING MULTIPLE
TITLE OF INVENTION: BINDING REGIONS DIRECTED TO TRANSCYTOTIC LIGANDS
FILE REFERENCE: 057220/1301
CURRENT APPLICATION NUMBER: US/09/949,039
CURRENT FILING DATE: 2001-09-06
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 37
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Gly-Ser linker
US-09-949-039-37

Query Match 100.0%; Score 5; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|
|
|
|
Db 1 GGGGS 5

RESULT 51

US-10-071-866-42
Sequence 42, Application US/10071866
Publication No. US20030165988A1
GENERAL INFORMATION:
APPLICANT: Hua, Shao-bing
APPLICANT: Pauling, Michelle H.
APPLICANT: Zhu, Li
TITLE OF INVENTION: HIGH THROUGHPUT GENERATION OF HUMAN MONOCLONAL ANTIBODY AGAINST P
TITLE OF INVENTION: FRAGMENTS DERIVED FROM MEMBRANE PROTEINS
FILE REFERENCE: 25636-717
CURRENT APPLICATION NUMBER: US/10/071,866
CURRENT FILING DATE: 2002-02-08
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PatentIn version 3.1
SEQ ID NO 42
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Linker Sequence

US-10-071-866-42

Query Match 100.0%; Score 5; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
|||||
Db 1 GGGGS 5

RESULT 52

US-10-245-415B-8
; Sequence 8, Application US/10245415B
; Publication No. US20030166141A1
; GENERAL INFORMATION:
; APPLICANT: Case, Casey Christopher
; APPLICANT: Cox III, George N.
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Liu, Qiang
; APPLICANT: Rebar, Edward J.
; TITLE OF INVENTION: REGULATION OF ENDOGENOUS GENE EXPRESSION IN CELLS
; TITLE OF INVENTION: USING ZINC FINGER PROTEINS
; FILE REFERENCE: 8325-0002.22 / S2-US7
; CURRENT APPLICATION NUMBER: US/10/245,415B
; CURRENT FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:linker

US-10-245-415B-8

Query Match 100.0%; Score 5; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
|||||
Db 1 GGGGS 5

RESULT 53

US-10-285-976-207
; Sequence 207, Application US/10285976
; Publication No. US20030165500A1
; GENERAL INFORMATION:
; APPLICANT: Rhee, Chae-Seo
; APPLICANT: Malini, Sen
; APPLICANT: Wu, Christina
; APPLICANT: Leoni, Lorenzo M.
; APPLICANT: Carr, Maripat
; APPLICANT: Carson, Dennis A.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Wnt and Frizzled Receptors as Targets for Immunotherapy
; TITLE OF INVENTION: in Head and Neck Squamous Cell Carcinomas
; FILE REFERENCE: 023070-130320US
; CURRENT APPLICATION NUMBER: US/10/285,976
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: US 60/287,995
; PRIOR FILING DATE: 2001-05-01
; PRIOR APPLICATION NUMBER: WO PCT/US02/13802
; PRIOR FILING DATE: 2002-05-01
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 207
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:peptide linker
US-10-285-976-207

Query Match 100.0%; Score 5; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
|||||
Db 1 GGGGS 5

RESULT 54

US-10-338-083-22
; Sequence 22, Application US/10338083
; Publication No. US20030166559A1
; GENERAL INFORMATION:
; APPLICANT: Desjarlais, John R.
; APPLICANT: Tansey, Malu Lourdes G.
; APPLICANT: Dahiyat, Bassil I.
; TITLE OF INVENTION: Dominant Negative Proteins and Methods Thereof
; FILE REFERENCE: A-71273-2
; CURRENT APPLICATION NUMBER: US/10/338,083
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US 60/345,805
; PRIOR FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: US 60/373,453
; PRIOR FILING DATE: 2002-04-17
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: linker consensus sequence

US-10-338-083-22

Query Match 100.0%; Score 5; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
|||||
Db 1 GGGGS 5

RESULT 55

US-10-247-960-1
; Sequence 1, Application US/10247960
; Publication No. US20030175304A1
; GENERAL INFORMATION:
; APPLICANT: Hawaii Biotechnology Group, Inc.
; APPLICANT: Peters, Iain
; APPLICANT: Coller, Beth-Ann
; APPLICANT: McDonnell, Michael
; APPLICANT: Ivy, John
; APPLICANT: Harada, Kent
; TITLE OF INVENTION: RECOMBINANT DIMERIC ENVELOPE VACCINE
; TITLE OF INVENTION: AGAINST FLAVIVIRAL INFECTION
; FILE REFERENCE: 24733-20005.01
; CURRENT APPLICATION NUMBER: US/10/247,960
; CURRENT FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: 09/376,463
; PRIOR FILING DATE: 1999-08-18
; PRIOR APPLICATION NUMBER: 08/904,227
; PRIOR FILING DATE: 1997-07-31
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

FEATURE:
OTHER INFORMATION: Linker
JS-10-247-960-1

Query Match 100.0%; Score 5; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
|
|
|
|
Db 1 GGGGS 5

RESULT 56
JS-10-309-587A-1
Sequence 1, Application US/10309587A
Publication No. US20030175887A1
GENERAL INFORMATION:
APPLICANT: SHORT, JAY M.
TITLE OF INVENTION: SATURATION MUTAGENESIS IN DIRECTED EVOLUTION
FILE REFERENCE: DIV-1460-16A
CURRENT APPLICATION NUMBER: US/10/309,587A
CURRENT FILING DATE: 2002-12-04
PRIOR APPLICATION NUMBER: 09/185,373
PRIOR FILING DATE: 1998-11-03
PRIOR APPLICATION NUMBER: 08/760,489
PRIOR FILING DATE: 1996-12-05
PRIOR APPLICATION NUMBER: 60/008,311
PRIOR FILING DATE: 1995-12-07
PRIOR APPLICATION NUMBER: 08/962,504
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 08/677,112
PRIOR FILING DATE: 1996-07-09
PRIOR APPLICATION NUMBER: 08/651,569
PRIOR FILING DATE: 1996-05-22
PRIOR APPLICATION NUMBER: 60/008,316
PRIOR FILING DATE: 1995-12-07
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: spacer peptide
JS-10-309-587A-1

Query Match 100.0%; Score 5; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
|
|
|
|
Db 1 GGGGS 5

RESULT 57
JS-10-412-105-45
Sequence 45, Application US/10412105
Publication No. US20030175790A1
GENERAL INFORMATION:
APPLICANT: Case, Casey
TITLE OF INVENTION: CELLS FOR DRUG DISCOVERY
FILE REFERENCE: 8325-0010
CURRENT APPLICATION NUMBER: US/10/412,105
CURRENT FILING DATE: 2003-04-10
PRIOR APPLICATION NUMBER: 09/779,233
PRIOR FILING DATE: 2001-02-08
NUMBER OF SEQ ID NOS: 45
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 45
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: linker
US-10-412-105-45

TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: linker
US-10-412-105-45

Query Match 100.0%; Score 5; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
|
|
|
|
Db 1 GGGGS 5

RESULT 58
US-10-214-446-57
Sequence 57, Application US/10214446
Publication No. US20030180742A1
GENERAL INFORMATION:
APPLICANT: Weiner, David
APPLICANT: Burk, Mark J.
APPLICANT: Hitchman, Tim
APPLICANT: Pujol, Catherine
APPLICANT: Richardson, Toby
APPLICANT: Short, Jay M.
TITLE OF INVENTION: P450 ENZYMES, NUCLEIC ACIDS ENCODING
TITLE OF INVENTION: THEM AND METHODS OF MAKING AND USING THEM
FILE REFERENCE: 09010-5000c1
CURRENT APPLICATION NUMBER: US/10/214,446
CURRENT FILING DATE: 2002-08-05
PRIOR APPLICATION NUMBER: US 60/309,497
PRIOR FILING DATE: 2001-08-03
NUMBER OF SEQ ID NOS: 59
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 57
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: linker peptide
US-10-214-446-57

Query Match 100.0%; Score 5; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
|
|
|
|
Db 1 GGGGS 5

RESULT 59
US-10-412-109-45
Sequence 45, Application US/10412109
Publication No. US20030180713A1
GENERAL INFORMATION:
APPLICANT: Case, Casey
TITLE OF INVENTION: CELLS FOR DRUG DISCOVERY
FILE REFERENCE: 8325-0010
CURRENT APPLICATION NUMBER: US/10/412,109
CURRENT FILING DATE: 2003-04-10
PRIOR APPLICATION NUMBER: US/09/779,233
PRIOR FILING DATE: 2001-02-08
NUMBER OF SEQ ID NOS: 45
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 45
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: linker
US-10-412-109-45

```
Query Match          100.0%; Score 5; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
DB 1 GGGGS 5

RESULT 60
US-10-222-026A-40
; Sequence 40, Application US/10222026A
; Publication No. US20030186322A1
; GENERAL INFORMATION:
; APPLICANT: Janda, Kim D.
; APPLICANT: Wirsching, Peter
; APPLICANT: Lerner, Richard A.
; APPLICANT: Gao, Changshou
; TITLE OF INVENTION: METHODS FOR DISPLAY OF HETERODIMERIC
; TITLE OF INVENTION: PROTEINS ON FILAMENTOUS PHAGE USING pVII AND pIX
; TITLE OF INVENTION: COMPOSITIONS, VECTORS AND COMBINATORIAL LIBRARIES
; FILE REFERENCE: TSRI 693.0 D1
; CURRENT APPLICATION NUMBER: US/10/222,026A
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: US 09/318,786
; PRIOR FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-10-222-026A-40

Query Match          100.0%; Score 5; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
DB 1 GGGGS 5

RESULT 61
US-10-246-229-64
; Sequence 64, Application US/10246229
; Publication No. US20030186356A1
; GENERAL INFORMATION:
; APPLICANT: Stemmer, Willem P.C.
; APPLICANT: Cramer, Andreas M.
; TITLE OF INVENTION: Methods for Generating Polynucleotides
; Having Desired Characteristics by Iterative Selection and
; Recombination
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/246,229
; FILING DATE: 17-Sep-2002
; CLASSIFICATION: <Unknown>
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/621,859
; FILING DATE: 25-MAR-1996
; APPLICATION NUMBER: US 08/564,955
; FILING DATE: 30-NOV-1995
; APPLICATION NUMBER: US 08/537,874
; FILING DATE: 30-OCT-1995
; APPLICATION NUMBER: WO PCT/US95/02126
; FILING DATE: 17-FEB-1995
; APPLICATION NUMBER: US 08/198,431
; FILING DATE: 17-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 018097-014612US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 64:
US-10-246-229-64

Query Match          100.0%; Score 5; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
DB 1 GGGGS 5

RESULT 62
US-10-099-816B-3
; Sequence 3, Application US/10099816B
; Publication No. US20030194763A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; TITLE OF INVENTION: End Selection in Directed Evolution
; FILE REFERENCE: 112766.224
; CURRENT APPLICATION NUMBER: US/10/099,816B
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: US 09/522,289
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: US 09/498,557
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 09/495,052
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: US 09/332,835
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/276,860
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: US 09/267,118
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: US 09/246,178
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: US 09/185,373
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: US 08/760,489
; PRIOR FILING DATE: 1996-12-05
; PRIOR APPLICATION NUMBER: US 60/008,311
; PRIOR FILING DATE: 1995-12-07
; Remaining prior application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 5
```

TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: spacer peptide sequence of any number
JS-10-099-816B-3

Query Match 100.0%; Score 5; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|||||
DB 1 GGGGS 5

RESULT 63

JS-10-087-426-3
Sequence 3, Application US/10087426
Publication No. US20020142394A1
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: SHORT, Jay M.
TITLE OF INVENTION: EXONUCLEASE-MEDIATED GENE ASSEMBLY IN DIRECTED EVOLUTION
FILE REFERENCE: DIVER1460-23
CURRENT APPLICATION NUMBER: US/10/087,426
CURRENT FILING DATE: 2002-03-01
PRIOR APPLICATION NUMBER: US 09/276,860
PRIOR FILING DATE: 1999-03-26
PRIOR APPLICATION NUMBER: US 09/267,118
PRIOR FILING DATE: 1999-03-09
PRIOR APPLICATION NUMBER: US 09/246,178
PRIOR FILING DATE: 1999-02-04
PRIOR APPLICATION NUMBER: US 09/185,373
PRIOR FILING DATE: 1998-11-03
PRIOR APPLICATION NUMBER: US 08/760,489
PRIOR FILING DATE: 1996-11-05
PRIOR APPLICATION NUMBER: US 60/008,311
PRIOR FILING DATE: 1995-11-07
PRIOR APPLICATION NUMBER: US 08/962,504
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: US 08/677,112
PRIOR FILING DATE: 1996-07-09
PRIOR APPLICATION NUMBER: US 08/651,568
PRIOR FILING DATE: 1996-05-22
PRIOR APPLICATION NUMBER: US 60/008,316
PRIOR FILING DATE: 1995-11-07
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Antibody spacer peptide. The entire peptide sequence can be
OTHER INFORMATION: repeated more than one time
JS-10-087-426-3

Query Match 100.0%; Score 5; DB 14; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|||||
DB 1 GGGGS 5

RESULT 64

JS-10-057-505-15
Sequence 15, Application US/10057505
Publication No. US20020164674A1
GENERAL INFORMATION:
APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
APPLICANT: AURORA BIOSCIENCES CORPORATION

APPLICANT: TSIEH, Roger
APPLICANT: HEIM, Roger
APPLICANT: CUBITT, Andrew
TITLE OF INVENTION: TANDEM FLUORESCENT PROTEIN CONSTRUCTS
FILE REFERENCE: REGEN1260-3
CURRENT APPLICATION NUMBER: US/10/057,505
CURRENT FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: US 08/792,553
PRIOR FILING DATE: 1997-01-31
PRIOR APPLICATION NUMBER: US 09/396,003
PRIOR FILING DATE: 1999-09-13
NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn version 3.1
SEQ ID NO 15
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Cleavage recognition sequence
US-10-057-505-15

Query Match 100.0%; Score 5; DB 14; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|||||
DB 1 GGGGS 5

RESULT 65

US-10-115-984-6
Sequence 6, Application US/10115984
Publication No. US20020168377A1
GENERAL INFORMATION:
APPLICANT: Dr. Sch.tzl, Herrmann
TITLE OF INVENTION: Prion protein dimers
FILE REFERENCE: S1044
CURRENT APPLICATION NUMBER: US/10/115,984
CURRENT FILING DATE: 2002-04-05
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:
OTHER INFORMATION: peptide spacer
US-10-115-984-6

Query Match 100.0%; Score 5; DB 14; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|||||
DB 1 GGGGS 5

RESULT 66

US-10-153-159-18
Sequence 18, Application US/10153159
Publication No. US20020177170A1
GENERAL INFORMATION:
APPLICANT: Luo, Peter
APPLICANT: Hsieh, Mark
APPLICANT: Zhong, Pingyu
APPLICANT: Wang, Gaili
TITLE OF INVENTION: STRUCTURE-BASED SELECTION AND AFFINITY MATURATION OF ANTIBODY LIB
FILE REFERENCE: 26050-704
CURRENT APPLICATION NUMBER: US/10/153,159

; CURRENT FILING DATE: 2002-05-20
 ; PRIOR APPLICATION NUMBER: US 10/125,687
 ; PRIOR FILING DATE: 2002-04-17
 ; PRIOR APPLICATION NUMBER: US 60/284,407
 ; PRIOR FILING DATE: 2001-04-17
 ; NUMBER OF SEQ ID NOS: 125
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 18
 ; LENGTH: 5
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Flexon
 US-10-153-159-18

Query Match 100.0%; Score 5; DB 14; Length 5;
 Best Local Similarity 100.0%; Pred. No. 5.7e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
 |||||
 Db 1 GGGGS 5

RESULT 67
 US-10-006-069A-212
 ; Sequence 212, Application US/10006069A
 ; Publication No. US20030021776A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rebar, Edward
 ; APPLICANT: Jamieson, Andrew
 ; APPLICANT: Liu, Qiang
 ; APPLICANT: Liu, Pei-Qi
 ; APPLICANT: Wolfe, Alan
 ; APPLICANT: Eisenberg, Stephen P.
 ; APPLICANT: Jarvis, Eric
 ; APPLICANT: Sangamo Biosciences, Inc.
 ; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
 ; FILE REFERENCE: 019496-005830US
 ; CURRENT APPLICATION NUMBER: US/10/006,069A
 ; CURRENT FILING DATE: 2001-12-17
 ; PRIOR APPLICATION NUMBER: US 09/733,604
 ; PRIOR FILING DATE: 2000-12-07
 ; PRIOR APPLICATION NUMBER: US 09/736,083
 ; PRIOR FILING DATE: 2000-12-12
 ; PRIOR APPLICATION NUMBER: US 09/846,033
 ; PRIOR FILING DATE: 2001-04-30
 ; NUMBER OF SEQ ID NOS: 252
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 212
 ; LENGTH: 5
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: peptide linker
 US-10-006-069A-212

Query Match 100.0%; Score 5; DB 15; Length 5;
 Best Local Similarity 100.0%; Pred. No. 5.7e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
 |||||
 Db 1 GGGGS 5

RESULT 68
 US-10-153-176-18
 ; Sequence 18, Application US/10153176
 ; Publication No. US2003002240A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Luo, Peter

; APPLICANT: Hsieh, Mark
 ; APPLICANT: Zhong, Pingyu
 ; APPLICANT: Wang, Caill
 ; APPLICANT: Cao, Yicheng
 ; APPLICANT: Li, Shengfeng
 ; APPLICANT: Liu, Shengjiang
 ; TITLE OF INVENTION: GENERATION AND AFFINITY MATURATION OF ANTIBODY LIBRARY IN SILLI
 ; FILE REFERENCE: 26050-701
 ; CURRENT APPLICATION NUMBER: US/10/153,176
 ; CURRENT FILING DATE: 2002-05-20
 ; PRIOR APPLICATION NUMBER: US 10/125,687
 ; PRIOR FILING DATE: 2002-04-17
 ; PRIOR APPLICATION NUMBER: US 60/284,407
 ; PRIOR FILING DATE: 2001-04-17
 ; NUMBER OF SEQ ID NOS: 125
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 18
 ; LENGTH: 5
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Flexon
 US-10-153-176-18

Query Match 100.0%; Score 5; DB 15; Length 5;
 Best Local Similarity 100.0%; Pred. No. 5.7e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
 |||||
 Db 1 GGGGS 5

RESULT 69
 US-10-108-077-3
 ; Sequence 3, Application US/10108077
 ; Publication No. US20030036116A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DIVERSA CORPORATION
 ; APPLICANT: SHORT, Jay
 ; APPLICANT: DJAVAKHSHVILI, Tsotne
 ; APPLICANT: Frey, Gerhard
 ; TITLE OF INVENTION: EXONUCLEASE-MEDIATED NUCLEIC ACID REASSEMBLY IN DIRECTED EVOLU
 ; FILE REFERENCE: DIVER1460-14
 ; CURRENT APPLICATION NUMBER: US/10/108,077
 ; CURRENT FILING DATE: 2002-03-26
 ; PRIOR APPLICATION NUMBER: US/09/535,754
 ; PRIOR FILING DATE: 2000-03-27
 ; PRIOR APPLICATION NUMBER: US 09/522,289
 ; PRIOR FILING DATE: 2000-03-09
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 3
 ; LENGTH: 5
 ; TYPE: PRT
 ; ORGANISM: Artificial sequence
 ; FEATURE:
 ; OTHER INFORMATION: Antibody spacer peptide. The entire peptide sequence can be r
 ; OTHER INFORMATION: ated more than one time
 US-10-108-077-3

Query Match 100.0%; Score 5; DB 15; Length 5;
 Best Local Similarity 100.0%; Pred. No. 5.7e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
 |||||
 Db 1 GGGGS 5

RESULT 70
 US-10-158-238-6
 ; Sequence 6, Application US/10158238

Publication No. US20030040604A1

GENERAL INFORMATION: Immunex Corporation

APPLICANT: Immunex Corporation

APPLICANT: Dirksen, John S.

APPLICANT: Marken, John S.

TITLE OF INVENTION: SIGLEC-12 POLYPEPTIDES, POLYNUCLEOTIDES, AND METHODS OF USE THERE

FILE REFERENCE: 3290-A

CURRENT APPLICATION NUMBER: US/10/158,238

CURRENT FILING DATE: 2002-05-29

PRIOR APPLICATION NUMBER: US 60/294,199

PRIOR FILING DATE: 2001-05-29

NUMBER OF SEQ ID NOS: 30

SOFTWARE: PatentIn version 3.1

SEQ ID NO 6

LENGTH: 5

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Peptide Linker sequence

S-10-158-238-6

Query Match 100.0%; Score 5; DB 15; Length 5;

Best Local Similarity 100.0%; Pred. No. 5.7e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGGGS 5

b 1 GGGGS 5

RESULT 71

US-10-057-890A-16

Sequence 16, Application US/10057890A

Publication No. US20030044901A1

GENERAL INFORMATION:

APPLICANT: Coleman, Timothy

APPLICANT: Mansfield, Brian

TITLE OF INVENTION: Scaffold Fusion Polypeptides, Composition for Making the Same, an

TITLE OF INVENTION: Of Using the Same.

FILE REFERENCE: PF537

CURRENT APPLICATION NUMBER: US/10/057,890A

CURRENT FILING DATE: 2002-01-29

PRIOR APPLICATION NUMBER: 60/265,782

PRIOR FILING DATE: 2001-01-31

PRIOR APPLICATION NUMBER: 60/265,858

PRIOR FILING DATE: 2001-01-31

NUMBER OF SEQ ID NOS: 32

SEQ ID NO 16

LENGTH: 5

TYPE: PRT

ORGANISM: Homo sapiens

S-10-057-890A-16

Query Match 100.0%; Score 5; DB 15; Length 5;

Best Local Similarity 100.0%; Pred. No. 5.7e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGGGS 5

b 1 GGGGS 5

RESULT 72

S-10-055-713-6

Sequence 6, Application US/10055713

Publication No. US20030044957A1

GENERAL INFORMATION:

APPLICANT: JAMIESON, Andrew

APPLICANT: Li, Guofu

TITLE OF INVENTION: ZINC FINGER PROTEINS FOR DNA BINDING AND GENE

TITLE OF INVENTION: REGULATION IN PLANTS

FILE REFERENCE: 8325-0026 / S26-US1

CURRENT APPLICATION NUMBER: US/10/055,713

; CURRENT FILING DATE: 2002-06-17

; PRIOR APPLICATION NUMBER: 60/263,445

; PRIOR FILING DATE: 2001-01-22

; PRIOR APPLICATION NUMBER: 60/290,716

; PRIOR FILING DATE: 2001-05-11

; NUMBER OF SEQ ID NOS: 105

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 6

; LENGTH: 5

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: peptide ZFP linker

US-10-055-713-6

Query Match 100.0%; Score 5; DB 15; Length 5;

Best Local Similarity 100.0%; Pred. No. 5.7e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGGGS 5

Db 1 GGGGS 5

RESULT 73

US-10-045-465-1

; Sequence 1, Application US/10045465

; Publication No. US20030049605A1

; GENERAL INFORMATION:

; APPLICANT: van Es, Helmut

; TITLE OF INVENTION: Display of viral proteins

; FILE REFERENCE: 2183-4079US

; CURRENT APPLICATION NUMBER: US/10/045,465

; CURRENT FILING DATE: 2001-11-07

; PRIOR APPLICATION NUMBER: 09/315,244

; PRIOR FILING DATE: 1999-05-20

; PRIOR APPLICATION NUMBER: EP 98201678.4

; PRIOR FILING DATE: 1998-05-20

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 1

; LENGTH: 5

; TYPE: PRT

; ORGANISM: Moloney murine leukemia virus

; FEATURE:

; NAME/KEY: PEPTIDE

; LOCATION: (1) - (5)

; OTHER INFORMATION: /No. US20030049605A1e="linker sequence"

US-10-045-465-1

Query Match 100.0%; Score 5; DB 15; Length 5;

Best Local Similarity 100.0%; Pred. No. 5.7e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGGGS 5

Db 1 GGGGS 5

RESULT 74

US-10-127-427-5

; Sequence 5, Application US/10127427

; Publication No. US20030051275A1

; GENERAL INFORMATION:

; APPLICANT: Paul CHRISTOU; Eva STROGER; Rainer FISCHER; Carmen MARTIN-VAQUERO;

; APPLICANT: Stefan SC

; TITLE OF INVENTION: METHODS AND MEANS FOR EXPRESSION OF MAMMALIAN

; POLYPEPTIDES

; NUMBER OF SEQUENCES: 43

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fulbright & Jaworski L.L.P.

; STREET: 666 Fifth Avenue

; CITY: New York City

STATE: New York
COUNTRY: USA
ZIP: 10103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.25 inch, 1.44mb
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/127,427
FILING DATE: 23-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/333,527
FILING DATE: <Unknown>
APPLICATION NUMBER: 60/089,322
FILING DATE: June 15, 1998
ATTORNEY/AGENT INFORMATION:
NAME: Mary Anne Schofield
REGISTRATION NUMBER: 36,669
REFERENCE/DOCKET NUMBER: KL/JIC 202.1 - JEL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 318-3000
TELEFAX: (212) 752-5958
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 5
TYPE: amino acid
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-127-427-5

Query Match 100.0%; Score 5; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
Db 1 GGGGS 5

RESULT 75
US-10-145-280-4
; Sequence 4, Application US/10145280
; Publication No. US2003007620A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay M.
; APPLICANT: KELLER, Martin
; TITLE OF INVENTION: HIGH THROUGHPUT SCREENING FOR SEQUENCES OF INTEREST
; FILE REFERENCE: DIVER1280-20
; CURRENT APPLICATION NUMBER: US/10/145,280
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 09/685,432
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 09/444,112
; PRIOR FILING DATE: 1999-11-22
; PRIOR APPLICATION NUMBER: US 09/098,206
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: US 08/876,276
; PRIOR FILING DATE: 1997-06-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Linker peptide
US-10-145-280-4

Query Match 100.0%; Score 5; DB 15; Length 5;

Best Local Similarity 100.0%; Pred. No. 5.7e+05; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGGGS 5
Db 1 GGGGS 5
RESULT 76
US-10-112-788-14
; Sequence 14, Application US/10112788
; Publication No. US2003007676A1
; GENERAL INFORMATION:
; APPLICANT: DENARDO, SALLY
; APPLICANT: WINTROP, MICHELLE
; APPLICANT: DENARDO, GERALD
; TITLE OF INVENTION: ANTI-MUC-1 SINGLE CHAIN ANTIBODIES FOR TUMOR TARGETING
; FILE REFERENCE: 309T-000210US
; CURRENT APPLICATION NUMBER: US/10/112,788
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 60/280,721
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombinant peptide linker
US-10-112-788-14

Query Match 100.0%; Score 5; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
Db 1 GGGGS 5

RESULT 77
US-10-145-281-4
; Sequence 4, Application US/10145281
; Publication No. US2003007677A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay M.
; APPLICANT: KELLER, Martin
; TITLE OF INVENTION: HIGH THROUGHPUT SCREENING FOR SEQUENCES OF INTEREST
; FILE REFERENCE: DIVER1280-21
; CURRENT APPLICATION NUMBER: US/10/145,281
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 09/685,432
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 09/444,112
; PRIOR FILING DATE: 1999-11-22
; PRIOR APPLICATION NUMBER: US 09/098,206
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: US 08/876,276
; PRIOR FILING DATE: 1997-06-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Linker peptide
US-10-145-281-4

Query Match 100.0%; Score 5; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGGGS 5
b 1 GGGGS 5

RESULT 78

S-10-057-408-23
Sequence 23, Application US/10057408
Publication No. US20030082561A1
GENERAL INFORMATION:
APPLICANT: Seta, Takashi
TITLE OF INVENTION: Zinc Finger Domain Recognition Code and Uses Thereof
FILE REFERENCE: 109845.135
CURRENT APPLICATION NUMBER: US/10/057,408
CURRENT FILING DATE: 2002-01-23
PRIOR APPLICATION NUMBER: US 60/220,060
PRIOR FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 69
SOFTWARE: PatentIn version 3.0
SEQ ID NO 23
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Flexible linker

US-10-057-408-23

Query Match 100.0%; Score 5; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGGGS 5
b 1 GGGGS 5

RESULT 79

S-10-133-128-240
Sequence 240, Application US/10133128
Publication No. US20030082630A1
GENERAL INFORMATION:
APPLICANT: KOLKMAN, JOOST A.
TITLE OF INVENTION: COMBINATORIAL LIBRARIES OF MONOMER DOMAINS
FILE REFERENCE: 0319.410US
CURRENT APPLICATION NUMBER: US/10/133,128
CURRENT FILING DATE: 2002-04-26
PRIOR APPLICATION NUMBER: 60/374,107
PRIOR FILING DATE: 2002-04-18
PRIOR APPLICATION NUMBER: 60/333,359
PRIOR FILING DATE: 2001-11-26
PRIOR APPLICATION NUMBER: 60/337,209
PRIOR FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/286,823
PRIOR FILING DATE: 2001-04-26
NUMBER OF SEQ ID NOS: 244
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 240
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: peptide linker

US-10-133-128-240

Query Match 100.0%; Score 5; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGGGS 5

Db 1 GGGGS 5

RESULT 80
US-10-252-408-22
Sequence 22, Application US/10252408
Publication No. US20030082736A1
GENERAL INFORMATION:
APPLICANT: SMITH, Craig A.
TITLE OF INVENTION: TUMOR NECROSIS FACTOR-ALPHA AND BETA-RECEPTORS
FILE REFERENCE: A-71592
CURRENT APPLICATION NUMBER: US/10/252,408
CURRENT FILING DATE: 2002-09-24
PRIOR APPLICATION NUMBER: US/08/406,824
PRIOR FILING DATE: 1995-03-20
PRIOR APPLICATION NUMBER: US 08/255,849
PRIOR FILING DATE: 1994-06-08
PRIOR APPLICATION NUMBER: US 07/860,710
PRIOR FILING DATE: 1992-03-30
PRIOR APPLICATION NUMBER: US 07/523,635
PRIOR FILING DATE: 1990-05-10
PRIOR APPLICATION NUMBER: US 07/421,417
PRIOR FILING DATE: 1989-10-13
PRIOR APPLICATION NUMBER: US 07/405,370
PRIOR FILING DATE: 1989-09-11
PRIOR APPLICATION NUMBER: US 07/403,241
PRIOR FILING DATE: 1989-09-05
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn version 3.1
SEQ ID NO 22
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Miscellaneous Structure

US-10-252-408-22

Query Match 100.0%; Score 5; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 1 GGGGS 5

RESULT 81
US-10-196-793A-50
Sequence 50, Application US/10196793A
Publication No. US20030086919A1
GENERAL INFORMATION:
APPLICANT: ROSENBLUM, MARTIN G.
TITLE OF INVENTION: THERAPEUTIC AGENTS COMPRISING PRO-APOPTOTIC PROTEINS
FILE REFERENCE: CLFR:012US
CURRENT APPLICATION NUMBER: US/10/196,793A
CURRENT FILING DATE: 2002-07-17
PRIOR APPLICATION NUMBER: 60/360,361
PRIOR FILING DATE: 2002-02-28
PRIOR APPLICATION NUMBER: 60/332,886
PRIOR FILING DATE: 2001-11-06
PRIOR APPLICATION NUMBER: 60/306,091
PRIOR FILING DATE: 2001-07-17
NUMBER OF SEQ ID NOS: 61
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 50
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Linker

US-10-196-793A-50

Query Match 100.0%; Score 5; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|
|
|
|
|
Db 1 GGGGS 5

RESULT 82
US-10-247-279-1
; Sequence 1, Application US/10247279
; Publication No. US20030086928A1
; GENERAL INFORMATION:
; APPLICANT: Montminy, Marc
; APPLICANT: Wagner, Brandee
; TITLE OF INVENTION: COMPOSITIONS ASSOCIATED WITH COMPLEX
; FILE REFERENCE: FORMATION AND METHODS OF USE THEREOF
; CURRENT APPLICATION NUMBER: US/10/247,279
; CURRENT FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 60/190,705
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: PCTUS 01/08946
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Linker
US-10-247-279-1

Query Match 100.0%; Score 5; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|
|
|
|
|
Db 1 GGGGS 5

RESULT 83
US-10-014-318-1
; Sequence 1, Application US/10014318
; Publication No. US20030091986A1
; GENERAL INFORMATION:
; APPLICANT: Pallavicini, Maria G.
; APPLICANT: Mullaney, Brian P.
; TITLE OF INVENTION: Identification of Expressed Genes Using Phage Display
; FILE REFERENCE: 023070-120500US
; CURRENT APPLICATION NUMBER: US/10/014,318
; CURRENT FILING DATE: 2002-05-31
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: spacer peptide
; NAME/KEY: REPEAT
; LOCATION: (1)..(5)
; OTHER INFORMATION: positions 1-5 may be repeated an undefined number
; OTHER INFORMATION: of times
US-10-014-318-1

Query Match 100.0%; Score 5; DB 15; Length 5;

Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|
|
|
|
|
Db 1 GGGGS 5

RESULT 84
US-10-145-314-4
; Sequence 4, Application US/10145314
; Publication No. US20030092022A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay M.
; APPLICANT: KELLER, Martin
; TITLE OF INVENTION: HIGH THROUGHPUT SCREENING FOR SEQUENCES OF INTEREST
; FILE REFERENCE: DIVER1280-19
; CURRENT APPLICATION NUMBER: US/10/145,314
; CURRENT FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: US 05/685,432
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 09/444,112
; PRIOR FILING DATE: 1999-11-22
; PRIOR APPLICATION NUMBER: US 09/098,206
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: US 08/876,276
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Linker peptide
US-10-145-314-4

Query Match 100.0%; Score 5; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|
|
|
|
|
Db 1 GGGGS 5

RESULT 85
US-10-013-173-47
; Sequence 47, Application US/10013173
; Publication No. US20030095977A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen C.
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James A.
; APPLICANT: Reno, Jonh M.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; FILE REFERENCE: 690022.547C1
; CURRENT APPLICATION NUMBER: US/10/013,173
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Gly Ser linker
US-10-013-173-47

```
Query Match      100.0%; Score 5; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 1 GGGGS 5
  |||||
b 1 GGGGS 5

RESULT 86
S-10-141-531-206
Sequence 206, Application US/10141531
Publication No. US20030100743A1
GENERAL INFORMATION:
APPLICANT: Briggs, Steven P.
APPLICANT: Dalma, Bipin K.
APPLICANT: del Val, Greg
APPLICANT: Desjarlais, John R.
APPLICANT: Helfetz, Peter
APPLICANT: Luginbuhl, Peter
APPLICANT: Muchhal, Umesh
TITLE OF INVENTION: Nucleic Acids and Proteins with Thioresdoxin Reductase Activity
FILE REFERENCE: A-71457-2/RTY/RMS/RMK
CURRENT APPLICATION NUMBER: US/10/141,531
CURRENT FILING DATE: 2002-05-06
PRIOR APPLICATION NUMBER: US 60/370,609
PRIOR FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: US 60/289,029
PRIOR FILING DATE: 2001-05-04
NUMBER OF SEQ ID NOS: 208
SOFTWARE: PatentIn version 3.1
SEQ ID NO 206
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: linker consensus sequence

US-10-141-531-206

Query Match      100.0%; Score 5; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 1 GGGGS 5
  |||||
b 1 GGGGS 5

RESULT 87
S-10-150-762-47
Sequence 47, Application US/10150762
Publication No. US20030103948A1
GENERAL INFORMATION:
APPLICANT: Goshorn, Stephen C.
APPLICANT: Graves, Scott S.
APPLICANT: Schultz, Joanne E.
APPLICANT: Lin, Yukang
APPLICANT: Sanderson, James A.
APPLICANT: Reno, Jonh M.
APPLICANT: Dearslyne, Erica A.
TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
METHODS OF USE THEREOF
FILE REFERENCE: 690022.547C2
CURRENT APPLICATION NUMBER: US/10/150,762
CURRENT FILING DATE: 2002-05-17
NUMBER OF SEQ ID NOS: 90
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 47
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Gly Ser linker

US-10-150-762-47

Query Match      100.0%; Score 5; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 1 GGGGS 5
  |||||
b 1 GGGGS 5

RESULT 88
US-10-283-300-36
Sequence 36, Application US/10283300
Publication No. US20030103972A1
GENERAL INFORMATION:
APPLICANT: Memorial Sloan-Kettering Cancer Center
TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF MULTIPLE SPLICE VARIANTS
OF THE MU-OPIOID RECEPTOR GENE
FILE REFERENCE: 830002-2000.3
CURRENT APPLICATION NUMBER: US/10/283,300
CURRENT FILING DATE: 2002-10-29
PRIOR APPLICATION NUMBER: 09/761,962
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 09/743,872
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: PCT/US99/15974
PRIOR FILING DATE: 1999-07-15
NUMBER OF SEQ ID NOS: 46
SOFTWARE: PatentIn version 3.0
SEQ ID NO 36
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: basic unit of a linking peptide

US-10-283-300-36

Query Match      100.0%; Score 5; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 1 GGGGS 5
  |||||
b 1 GGGGS 5

RESULT 89
US-10-113-424-3
Sequence 3, Application US/10113424
Publication No. US20030105593A1
GENERAL INFORMATION:
APPLICANT: Eisenberg, Stephen P.
APPLICANT: Case, Casey C.
APPLICANT: Cox III, George N.
APPLICANT: Jamieson, Andrew
APPLICANT: Rebar, Edward J.
APPLICANT: Sangamo Biosciences, Inc.
TITLE OF INVENTION: Selection of Sites for Targeting by Zinc Finger
PROTEINS AND METHODS OF DESIGNING ZINC FINGER PROTEINS
TO BIND TO PRESELECTED SITES
FILE REFERENCE: 019496-001800US
CURRENT APPLICATION NUMBER: US/10/113,424
CURRENT FILING DATE: 2002-03-28
PRIOR APPLICATION NUMBER: US/09/229,007A
PRIOR FILING DATE: 1999-01-12
NUMBER OF SEQ ID NOS: 97
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
```

; OTHER INFORMATION: Description of Artificial Sequence:peptide linker
US-10-113-424-3

Query Match 100.0%; Score 5; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|||||

Db 1 GGGGS 5

RESULT 90
US-10-055-711-6
; Sequence 6, Application US/10055711
; Publication No. US20030108880A1
; GENERAL INFORMATION:
; APPLICANT: REBAR, Edward
; TITLE OF INVENTION: MODIFIED ZINC FINGER BINDING PROTEINS
; FILE REFERENCE: 8325-0025
; CURRENT APPLICATION NUMBER: US/10/055,711
; CURRENT FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: peptide linker
US-10-055-711-6

Query Match 100.0%; Score 5; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|||||

Db 1 GGGGS 5

RESULT 91
US-10-270-071-1
; Sequence 1, Application US/10270071
; Publication No. US20030113333A1
; GENERAL INFORMATION:
; APPLICANT: ROSSI, EDMUND
; APPLICANT: CHANG, CHIEN-HSING KEN
; TITLE OF INVENTION: AFFINITY ENHANCEMENT AGENTS
; FILE REFERENCE: 042418/0110
; CURRENT APPLICATION NUMBER: US/10/270,071
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: 60/328,835
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 60/341,881
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Linker peptide
US-10-270-071-1

Query Match 100.0%; Score 5; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|||||

Db 1 GGGGS 5

RESULT 92
US-10-218-131-1
; Sequence 1, Application US/10218131
; Publication No. US20030113759A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; TITLE OF INVENTION: METHOD OF DNA SHUFFLING WITH
; TITLE OF INVENTION: POLYNUCLEOTIDES PRODUCED BY BLOCKING OR INTERRUPTING A
; FILE REFERENCE: DIVER1220-2
; CURRENT APPLICATION NUMBER: US/10/218,131
; PRIOR FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: US/09/214,645
; PRIOR FILING DATE: 1999-09-27
; PRIOR APPLICATION NUMBER: PCT/US97/12239
; PRIOR FILING DATE: 1997-07-09
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: linker peptide
US-10-218-131-1

Query Match 100.0%; Score 5; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|||||

Db 1 GGGGS 5

RESULT 93
US-10-207-655-39
; Sequence 39, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 39
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Linker polypeptide
US-10-207-655-39

Query Match 100.0%; Score 5; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|||||

Db 1 GGGGS 5

RESULT 94
US-10-288-250-7
; Sequence 7, Application US/10288250
; Publication No. US20030121068A1
; GENERAL INFORMATION:

APPLICANT: Orchard, Paul
APPLICANT: Lewis, Victor
APPLICANT: Regents of the University of Minnesota
TITLE OF INVENTION: Vector Encoding Suicide and Marker Constructs
FILE REFERENCE: 600.527US1

CURRENT APPLICATION NUMBER: US/10/288,250
CURRENT FILING DATE: 2002-11-04
PRIOR APPLICATION NUMBER: US 60/369,507
PRIOR FILING DATE: 2002-04-03
PRIOR APPLICATION NUMBER: US 60/334,795
PRIOR FILING DATE: 2001-11-30
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 7

LENGTH: 5

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: A linker.

S-10-288-250-7

Query Match 100.0%; Score 5; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 1 GGGGS 5

|||||

b 1 GGGGS 5

RESULT 95

US-10-218-102-417

Sequence 417, Application US/10218102

Publication No. US20030130827A1

GENERAL INFORMATION:

APPLICANT: Bentzien, Joerg

APPLICANT: Dahiya, Basil I.

APPLICANT: Desjarlais, John R.

APPLICANT: Hayes, Robert J.

APPLICANT: Vielmetter, Jost

TITLE OF INVENTION: Protein Design Automation for Protein Libraries

FILE REFERENCE: A-67229-11/RFT/RMS/RMK
CURRENT APPLICATION NUMBER: US/10/218,102

CURRENT FILING DATE: 2002-08-12

PRIOR APPLICATION NUMBER: US 09/927,790

PRIOR FILING DATE: 2001-08-10

PRIOR APPLICATION NUMBER: US 60/311,545

PRIOR FILING DATE: 2001-09-10

PRIOR APPLICATION NUMBER: US 60/324,899

PRIOR FILING DATE: 2001-09-25

PRIOR APPLICATION NUMBER: US 60/351,937

PRIOR FILING DATE: 2002-01-25

PRIOR APPLICATION NUMBER: US 60/352,103

PRIOR FILING DATE: 2002-01-25

NUMBER OF SEQ ID NOS: 432

SOFTWARE: PatentIn version 3.1

SEQ ID NO 417

LENGTH: 5

TYPE: PRT

ORGANISM: Artificial sequence

FEATURE:

OTHER INFORMATION: linker

US-10-218-102-417

Query Match 100.0%; Score 5; DB 16; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

zy 1 GGGGS 5

|||||

zb 1 GGGGS 5

RESULT 96

US-09-782-980-124

Sequence 124, Application US/09782980

Patent No. US20020072089A1

GENERAL INFORMATION:

APPLICANT: Khodadoust, Mehran M.

APPLICANT: MacBeth, Kyle J.

APPLICANT: Busfield, Samantha J.

APPLICANT: McCarthy, Sean A.

APPLICANT: Holtzman, Douglas A.

APPLICANT: Gu, Wei

APPLICANT: White, David

APPLICANT: Par, Yang

TITLE OF INVENTION: NOVEL ITALY, LOR-2, STRIFE, TRASH, BDSF, LRSG, AND

TITLE OF INVENTION: STMTS PROTEIN AND NUCLEIC ACID MOLECULES AND USES

TITLE OF INVENTION: THEREFOR

FILE REFERENCE: MNI-121CP

CURRENT APPLICATION NUMBER: US/09/782,980

CURRENT FILING DATE: 2001-02-13

PRIOR APPLICATION NUMBER: PC*/US00/02125

PRIOR FILING DATE: 2000-01-27

PRIOR APPLICATION NUMBER: 09/448,076

PRIOR FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: 09/276,400

PRIOR FILING DATE: 1999-03-25

PRIOR APPLICATION NUMBER: 60/117,580

PRIOR FILING DATE: 1999-01-27

PRIOR APPLICATION NUMBER: 09/014,195

PRIOR FILING DATE: 1998-01-27

PRIOR APPLICATION NUMBER: 09/014,348

PRIOR FILING DATE: 1998-01-27

PRIOR APPLICATION NUMBER: 09/086,892

PRIOR FILING DATE: 1998-05-29

PRIOR APPLICATION NUMBER: 09/296,208

PRIOR FILING DATE: 1999-04-21

PRIOR APPLICATION NUMBER: 09/063,950

PRIOR FILING DATE: 1998-04-21

PRIOR APPLICATION NUMBER: 09/561,381

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: 09/561,810

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: 09/087,121

PRIOR FILING DATE: 1998-05-29

PRIOR APPLICATION NUMBER: 09/672,721

PRIOR FILING DATE: 2000-09-28

PRIOR APPLICATION NUMBER: 09/049,799

PRIOR FILING DATE: 1998-03-27

NUMBER OF SEQ ID NOS: 176

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 124

LENGTH: 6

TYPE: PRT

ORGANISM: Homo sapiens

US-09-782-980-124

Query Match 100.0%; Score 5; DB 9; Length 6;

Best Local Similarity 100.0%; Pred. No. 5.7e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5

|||||

Db 1 GGGGS 5

RESULT 97

US-09-883-777-11

Sequence 11, Application US/09883777

Patent No. US20020110853A1

GENERAL INFORMATION:

APPLICANT: Wiley, Steven R.

TITLE OF INVENTION: TWEAK RECEPTOR

FILE REFERENCE: 2968-C

CURRENT APPLICATION NUMBER: US/09/883,777


```
; CURRENT FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/172,878
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: US 60/203,347
; PRIOR FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: PCT/US00/34755
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/742,454
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Linker Moiety
; NAME/KEY: misc.feature
; LOCATION: (6)..(6)
; OTHER INFORMATION: Xaa is one or more repeats of GGGGS
US-09-883-777-11

Query Match          100.0%; Score 5; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 1 GGGGS 5

RESULT 98
US-09-949-039-48
; Sequence 48, Application US/09949039
; Publication No. US20030166160A1
; GENERAL INFORMATION:
; APPLICANT: HAWLEY, STEPHEN B.
; TITLE OF INVENTION: COMPOUNDS AND MOLECULAR COMPLEXES COMPRISING MULTIPLE
; FILE REFERENCE: 057220/1301
; CURRENT APPLICATION NUMBER: US/09/949,039
; CURRENT FILING DATE: 2001-09-06
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide
US-09-949-039-48

Query Match          100.0%; Score 5; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 1 GGGGS 5

RESULT 99
US-10-158-238-7
; Sequence 7, Application US/10158238
; Publication No. US20030040604A1
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Dirk, Anderson M.
; APPLICANT: Marken, John S.
; TITLE OF INVENTION: SIGLEC-12 POLYPEPTIDES, POLYNUCLEOTIDES, AND METHODS OF USE THERE
; FILE REFERENCE: 3290-A
; CURRENT APPLICATION NUMBER: US/10/158,238
; CURRENT FILING DATE: 2002-05-29
```

```
; PRIOR APPLICATION NUMBER: US 60/294,199
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Linker sequence
; NAME/KEY: MISC FEATURE
; LOCATION: (6)..(6)
; OTHER INFORMATION: X at position 6 is one or more repeats of GGGGS
US-10-158-238-7

Query Match          100.0%; Score 5; DB 15; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 1 GGGGS 5

RESULT 100
US-09-815-837-112
; Sequence 112, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
; APPLICANT: Carter, Darrick
; APPLICANT: Zhu, Shirley
; APPLICANT: Arimilli, Subhashini
; APPLICANT: Wang, Aijun
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Immune Mediators and Related Methods
; FILE REFERENCE: 014058-00567005
; CURRENT APPLICATION NUMBER: US/09/815,837
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 60/191,274
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: US 60/204,249
; PRIOR FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: US 60/264,003
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 112
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:downstream
; OTHER INFORMATION: linker for C0583 and C0590
US-09-815-837-112

Query Match          100.0%; Score 5; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 3 GGGGS 7

Search completed: October 21, 2003, 18:49:43
Job time : 33.7222 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:27:21 ; Search time 13.8889 Seconds
(without alignments)
15.232 Million cell updates/sec

Title: US-10-057-890A-16

Perfect score: 5

Sequence: 1 GGGGS 5

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 423:0858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 90058

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Listing first 150 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTJUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5	100.0	5	1	US-07-959-946-6
2	5	100.0	5	1	US-08-176-500-140
3	5	100.0	5	1	US-08-471-052A-140
4	5	100.0	5	1	US-08-225-224-54
5	5	100.0	5	1	US-08-236-918A-18
6	5	100.0	5	1	US-08-463-163-1
7	5	100.0	5	1	US-08-566-800A-58
8	5	100.0	5	1	US-08-244-489-5
9	5	100.0	5	1	US-08-189-331-140
10	5	100.0	5	1	US-08-333-577-6
11	5	100.0	5	1	US-08-575-361A-32
12	5	100.0	5	2	US-08-564-955-64
13	5	100.0	5	2	US-08-528-523-13
14	5	100.0	5	2	US-08-537-874-62
15	5	100.0	5	2	US-08-448-418-86
16	5	100.0	5	2	US-08-471-939-140
17	5	100.0	5	2	US-08-471-800-140
18	5	100.0	5	2	US-08-932-589-58
19	5	100.0	5	2	US-08-647-449-26
20	5	100.0	5	2	US-08-821-840-1
21	5	100.0	5	2	US-08-902-623-75
22	5	100.0	5	2	US-08-308-494A-1
23	5	100.0	5	2	US-08-665-202-25
24	5	100.0	5	2	US-08-809-668-12
25	5	100.0	5	2	US-08-792-553-15
26	5	100.0	5	2	US-08-751-767A-24
27	5	100.0	5	2	US-08-478-373-144

28	5	100.0	5	2	US-08-818-253-40	Sequence 40, Appl
29	5	100.0	5	3	US-08-722-258-54	Sequence 54, Appl
30	5	100.0	5	3	US-09-397-951-12	Sequence 12, Appl
31	5	100.0	5	3	US-08-776-271-3	Sequence 3, Appl
32	5	100.0	5	3	US-08-621-859-64	Sequence 62, Appl
33	5	100.0	5	3	US-09-100-856A-62	Sequence 62, Appl
34	5	100.0	5	3	US-09-046-993-5	Sequence 5, Appl
35	5	100.0	5	3	US-09-215-035-3	Sequence 3, Appl
36	5	100.0	5	3	US-09-075-511-64	Sequence 64, Appl
37	5	100.0	5	3	US-09-099-015-64	Sequence 64, Appl
38	5	100.0	5	3	US-08-818-252-40	Sequence 40, Appl
39	5	100.0	5	3	US-09-091-814-1	Sequence 1, Appl
40	5	100.0	5	3	US-09-281-792B-26	Sequence 26, Appl
41	5	100.0	5	3	US-09-232-863-62	Sequence 62, Appl
42	5	100.0	5	3	US-09-362-805-1	Sequence 1, Appl
43	5	100.0	5	3	US-09-362-805-2	Sequence 2, Appl
44	5	100.0	5	3	US-09-367-953B-26	Sequence 26, Appl
45	5	100.0	5	3	US-09-133-508A-62	Sequence 62, Appl
46	5	100.0	5	3	US-09-165-060-64	Sequence 64, Appl
47	5	100.0	5	3	US-09-501-698-62	Sequence 62, Appl
48	5	100.0	5	3	US-09-173-190-1	Sequence 1, Appl
49	5	100.0	5	3	US-09-173-190-2	Sequence 2, Appl
50	5	100.0	5	4	US-09-240-310-64	Sequence 64, Appl
51	5	100.0	5	4	US-09-590-778-64	Sequence 64, Appl
52	5	100.0	5	4	US-09-150-864A-18	Sequence 18, Appl
53	5	100.0	5	4	US-09-535-754-3	Sequence 3, Appl
54	5	100.0	5	4	US-09-477-763-62	Sequence 62, Appl
55	5	100.0	5	4	US-09-215-212-1	Sequence 1, Appl
56	5	100.0	5	4	US-09-590-774-64	Sequence 64, Appl
57	5	100.0	5	4	US-09-619-550-64	Sequence 64, Appl
58	5	100.0	5	4	US-09-240-179-52	Sequence 52, Appl
59	5	100.0	5	4	US-09-602-373A-76	Sequence 76, Appl
60	5	100.0	5	4	US-09-240-734-64	Sequence 64, Appl
61	5	100.0	5	4	US-09-231-253-62	Sequence 62, Appl
62	5	100.0	5	4	US-09-006-428A-23	Sequence 23, Appl
63	5	100.0	5	4	US-09-724-958-62	Sequence 62, Appl
64	5	100.0	5	4	US-09-229-007A-3	Sequence 3, Appl
65	5	100.0	5	4	US-09-300-008B-64	Sequence 64, Appl
66	5	100.0	5	4	US-09-470-191-95	Sequence 95, Appl
67	5	100.0	5	4	US-09-316-919-56	Sequence 56, Appl
68	5	100.0	5	4	US-09-146-979-86	Sequence 86, Appl
69	5	100.0	5	4	US-08-840-713-43	Sequence 43, Appl
70	5	100.0	5	4	US-09-761-962A-36	Sequence 36, Appl
71	5	100.0	5	4	US-09-731-558-6	Sequence 6, Appl
72	5	100.0	5	4	US-09-696-322-64	Sequence 64, Appl
73	5	100.0	5	4	US-09-717-391-64	Sequence 64, Appl
74	5	100.0	5	4	US-09-315-574-25	Sequence 25, Appl
75	5	100.0	5	4	US-08-913-370-1	Sequence 1, Appl
76	5	100.0	5	4	US-09-696-313-64	Sequence 64, Appl
77	5	100.0	5	4	US-09-077-028A-17	Sequence 17, Appl
78	5	100.0	5	4	US-09-229-037-8	Sequence 8, Appl
79	5	100.0	5	4	US-08-406-824A-22	Sequence 22, Appl
80	5	100.0	5	4	US-09-223-040-8	Sequence 8, Appl
81	5	100.0	5	4	US-09-240-307-64	Sequence 64, Appl
82	5	100.0	5	4	US-09-640-198D-34	Sequence 34, Appl
83	5	100.0	5	5	PT-US92-0863A-6	Sequence 6, Appl
84	5	100.0	5	5	PT-US95-04468-54	Sequence 54, Appl
85	5	100.0	5	6	US-08-704-170-76	Sequence 76, Appl
86	5	100.0	5	6	US-08-809-668-10	Sequence 10, Appl
87	5	100.0	5	6	US-09-397-951-10	Sequence 10, Appl
88	5	100.0	5	6	PT-US94-02631-76	Sequence 76, Appl
89	5	100.0	5	7	US-08-902-623-49	Sequence 49, Appl
90	5	100.0	5	7	US-09-374-454-16	Sequence 16, Appl
91	5	100.0	5	7	US-08-897-956A-38	Sequence 38, Appl
92	5	100.0	5	8	US-08-750-128-6	Sequence 6, Appl
93	5	100.0	5	9	US-08-487-890A-144	Sequence 144, App
94	5	100.0	5	9	US-08-448-418-88	Sequence 88, Appl
95	5	100.0	5	9	US-08-478-435-144	Sequence 144, App
96	5	100.0	5	9	US-08-478-435-145	Sequence 145, App
97	5	100.0	5	9	US-08-337-483-144	Sequence 144, App
98	5	100.0	5	9	US-08-478-373-144	Sequence 144, App
99	5	100.0	5	9	US-08-478-373-144	Sequence 144, App
100	5	100.0	5	9	US-08-478-373-144	Sequence 144, App

101 Sequence 145, App
102 Sequence 144, App
103 Sequence 145, App
104 Sequence 144, App
105 Sequence 145, App
106 Sequence 144, App
107 Sequence 145, App
108 Sequence 144, App
109 Sequence 145, App
110 Sequence 144, App
111 Sequence 145, App
112 Sequence 144, App
113 Sequence 145, App
114 Sequence 144, App
115 Sequence 145, App
116 Sequence 144, App
117 Sequence 145, App
118 Sequence 144, App
119 Sequence 145, App
120 Sequence 144, App
121 Sequence 145, App
122 Sequence 144, App
123 Sequence 145, App
124 Sequence 144, App
125 Sequence 145, App
126 Sequence 144, App
127 Sequence 145, App
128 Sequence 144, App
129 Sequence 145, App
130 Sequence 144, App
131 Sequence 145, App
132 Sequence 144, App
133 Sequence 145, App
134 Sequence 144, App
135 Sequence 145, App
136 Sequence 144, App
137 Sequence 145, App
138 Sequence 144, App
139 Sequence 145, App
140 Sequence 144, App
141 Sequence 145, App
142 Sequence 144, App
143 Sequence 145, App
144 Sequence 144, App
145 Sequence 145, App
146 Sequence 144, App
147 Sequence 145, App
148 Sequence 144, App
149 Sequence 145, App
150 Sequence 144, App

ALIGNMENTS

RESULT 1
US-07-959-946-6
Sequence 6, Application US/07959946
Patent No. 5408038
GENERAL INFORMATION:
APPLICANT: Smith, Richard K.
APPLICANT: Koduri, Raju
APPLICANT: Young, Stephen G.
APPLICANT: Witstum, Joseph L.
APPLICANT: Curtiss, Linda K.
TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a
TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &
ADDRESSEE: Milnamow, Ltd.
STREET: 180 No. 5408038th Stetson, Suite 4700
CITY: Chicago

STATE: Illinois
COUNTRY: USA
ZIP: 60601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/959,946
FILING DATE: 19921008
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/901,706
FILING DATE: 18-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Gamson, Edward P.
REGISTRATION NUMBER: 29,381
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312)616-5400
TELEFAX: (312)616-5460
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-959-946-6
Query Match 100.0%; Score 5; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;
QY 1 GGGGS 5
DB 1 GGGGS 5
RESULT 2
US-08-176-500-140
Sequence 140, Application US/08176500
Patent No. 5498538
GENERAL INFORMATION:
APPLICANT: Kay, B. K.
APPLICANT: Fowkes, D. M.
TITLE OF INVENTION: Totally Synthetic Affinity Reagents
NUMBER OF SEQUENCES: 141
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/176,500
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/013,416
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Mirock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-143
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090

TELEFAX: 212 869-8864/9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 140:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 US-08-176-500-140

Query Match 100.0%; Score 5; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
 |||||
 Db 1 GGGGS 5

RESULT 3

US-08-471-052A-140
 Sequence 140, Application US/08471052A
 Patent No. 5625033

GENERAL INFORMATION:
 APPLICANT: Kay, B. K.
 TITLE OF INVENTION: Totally Synthetic Affinity Reagents
 NUMBER OF SEQUENCES: 166
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036-2711

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/471,052A
 FILING DATE: 06-JUNE-1995
 CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
 NAME: Mirock, S. Leslie
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 1101-179
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212 790-9090
 TELEFAX: 212 869-8864/9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 140:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 US-08-471-052A-140

Query Match 100.0%; Score 5; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
 |||||
 Db 1 GGGGS 5

RESULT 4

US-08-225-224-54

Sequence 54, Application US/08225224
 Patent No. 5635599
 GENERAL INFORMATION:
 APPLICANT: PASTAN, Ira
 TITLE OF INVENTION: CIRCULARLY PERMUTATED LIGANDS AND
 TITLE OF INVENTION: CIRCULARLY PERMUTATED FUSION PROTEINS
 NUMBER OF SEQUENCES: 57
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend Kourie and Crew
 STREET: Stewart Street Tower, One Market Plaza
 CITY: San Francisco
 STATE: California
 COUNTRY: US
 ZIP: 94105-1493

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/225,224
 FILING DATE: 8-APR-1994
 CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
 NAME: Weber, Ellen L.
 REGISTRATION NUMBER: 32,762
 REFERENCE/DOCKET NUMBER: 15280-193
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 543-9600
 TELEFAX: (415) 543-5043
 INFORMATION FOR SEQ ID NO: 54:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 US-08-225-224-54

Query Match 100.0%; Score 5; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
 |||||
 Db 1 GGGGS 5

RESULT 5

US-08-236-918A-18
 Sequence 18, Application US/08236918A
 Patent No. 5674704

GENERAL INFORMATION:
 APPLICANT: Alderson, Mark R.
 APPLICANT: Goodwin, Raymond G.
 APPLICANT: Smith, Craig A.
 TITLE OF INVENTION: Cytokine Designated 4-1BB Ligand
 NUMBER OF SEQUENCES: 18
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Kathryn A. Anderson, Immunex Corporation
 STREET: 51 University Street
 CITY: Seattle
 STATE: Washington
 COUNTRY: US
 ZIP: 98101

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: Apple Power Macintosh
 OPERATING SYSTEM: Apple 7.5.3
 SOFTWARE: Microsoft Word, Version #6.0.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/236,918A

FILING DATE: 06-May-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/060,843
FILING DATE: 07-May-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Anderson, Kathryn A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2801-B
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
S-08-236-918A-18

Query Match 100.0%; Score 5; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|||||
Db 1 GGGGS 5

RESULT 6
US-08-463-163-1
Sequence 1, Application US/08463163
Patent No. 5696237
GENERAL INFORMATION:
APPLICANT: Fitzgerald, David J.
APPLICANT: Chaudhary, Vijay K.
APPLICANT: Pastan, Ira H.
APPLICANT: Waldmann, Thomas A.
APPLICANT: Queen, Cary L.
TITLE OF INVENTION: Recombinant Antibody-Toxin Fusion Protein
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: One Market Plaza, Steuart Street Tower
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105-1492
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,163
FILING DATE: 05-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/227,227
FILING DATE: 22-JAN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/911,227
FILING DATE: 24-SEP-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,361
FILING DATE: 21-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/865,722
FILING DATE: 08-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen L.
REGISTRATION NUMBER: 32,762

REFERENCE/DOCKET NUMBER: 015280-12211
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-463-163-1

Query Match 100.0%; Score 5; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|||||
Db 1 GGGGS 5

RESULT 7
US-08-566-800A-58
Sequence 58, Application US/08566800A
Patent No. 5736364
GENERAL INFORMATION:
APPLICANT: Kelley, Robert F.
APPLICANT: Lazarus, Robert A.
APPLICANT: Lee, Geoffrey F.
TITLE OF INVENTION: No. 5736364e1 Factor VIIa Inhibitors
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/566,800A
FILING DATE: 04-Dec-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kubinec, Jeffrey S.
REGISTRATION NUMBER: 36,575
REFERENCE/DOCKET NUMBER: P0958B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-8228
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-566-800A-58

Query Match 100.0%; Score 5; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|||||
Db 1 GGGGS 5

RESULT 8

```

US-08-244-469-5
; Sequence 5, Application US/08244469
; Patent No. 5736387
; GENERAL INFORMATION:
; APPLICANT: Paul, Ralph W.
; APPLICANT: Overell, Robert
; TITLE OF INVENTION: ENVELOPE FUSION VECTORS FOR USE IN GENE
; DELIVERY
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/244,469
; FILING DATE: 01-JUN-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Dylan, Tyler M.
; REGISTRATION NUMBER: 37,612
; REFERENCE/DOCKET NUMBER: 22627-20007.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5660
; TELEFAX: (415) 494-0792
; TELEX: 706141 MRSNFOR SFO
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-244-469-5

Query Match 100.0%; Score 5; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
Db 1 GGGGS 5

RESULT 9
US-08-189-331-140
; Sequence 140, Application US/08189331
; Patent No. 5747334
; GENERAL INFORMATION:
; APPLICANT: Kay, B. K.
; APPLICANT: Fowlkes, D. M.
; TITLE OF INVENTION: Totally Synthetic Affinity Reagents
; NUMBER OF SEQUENCES: 186
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/189,331

```

```

; FILING DATE: Concurrently herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-155
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 140:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-189-331-140

Query Match 100.0%; Score 5; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
Db 1 GGGGS 5

RESULT 10
US-08-333-577-6
; Sequence 6, Application US/08333577
; Patent No. 5786206
; GENERAL INFORMATION:
; APPLICANT: Smith, Richard K.
; APPLICANT: Koduri, Raju
; APPLICANT: Young, Stephen G.
; APPLICANT: Witztum, Joseph L.
; APPLICANT: Curtiss, Linda K.
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &
; ADDRESSEE: Milnamow, Ltd.
; STREET: 180 No. 5786206th Stetson, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/333,577
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Gamson, Edward P.
; REGISTRATION NUMBER: 29,381
; REFERENCE/DOCKET NUMBER: SCRF 234.0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5400
; TELEFAX: (312) 616-5460
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-333-577-6

```

Query Match 100.0%; Score 5; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2y 1 GGGGS 5
 |||||
 Db 1 GGGGS 5

RESULT 11
 JS-08-575-361A-32
 Sequence 32, Application US/08575361A
 Patent No. 5792640
 GENERAL INFORMATION:
 APPLICANT: Chandrasegaran, Srinivasan
 TITLE OF INVENTION: A GENERAL METHOD TO CLONE HYBRID
 TITLE OF INVENTION: RESTRICTION ENDONUCLEASES USING lig GENE
 NUMBER OF SEQUENCES: 35
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Cushman Darby & Cushman L.L.P.
 STREET: 1100 New York Avenue, NW, Ninth Floor, East
 STREET: Tower
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20005-3918
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/575.361A
 FILING DATE: 20-DEC-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Kokulis, Paul N.
 REGISTRATION NUMBER: 16,773
 REFERENCE/DOCKET NUMBER: PNK/4130/213779/DJP
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-861-3000
 TELEFAX: 202-822-0944
 TELEX: 6714627 CUSH
 INFORMATION FOR SEQ ID NO: 32:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5 amino acids
 TYPE: amino acid
 TOPOLOGY: linear

JS-08-575-361A-32

Query Match 100.0%; Score 5; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2y 1 GGGGS 5
 |||||
 Db 1 GGGGS 5

RESULT 12
 JS-08-564-955-64
 Sequence 64, Application US/08564955
 Patent No. 5811238
 GENERAL INFORMATION:
 APPLICANT: STEMMER, WILLEM P.C.
 APPLICANT: CRAMERI, ANDREAS M.
 TITLE OF INVENTION: METHODS FOR GENERATING POLYNUCLEOTIDES
 TITLE OF INVENTION: HAVING DESIRED CHARACTERISTICS BY ITERATIVE SELECTION AND
 TITLE OF INVENTION: RECOMBINATION
 NUMBER OF SEQUENCES: 67
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: TOWNSEND AND TOWNSEND AND CREW
 STREET: TWO EMBARCADERO CENTER, 8TH FLOOR

CITY: SAN FRANCISCO
 STATE: CALIFORNIA
 COUNTRY: U.S.A.
 ZIP: 94111-3834
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/564.955
 FILING DATE: 30-NOV-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/198,431
 FILING DATE: 17-FEB-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/537,874
 FILING DATE: 30-OCT-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/02126
 FILING DATE: 17-FEB-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: DUNN, TRACY J.
 REGISTRATION NUMBER: 34,587
 REFERENCE/DOCKET NUMBER: 16528J-014611US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 326-2400
 TELEFAX: (415) 576-0300
 INFORMATION FOR SEQ ID NO: 64:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear

US-08-564-955-64

Query Match 100.0%; Score 5; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
 |||||
 Db 1 GGGGS 5

RESULT 13
 US-08-528-523-13
 Sequence 13, Application US/08528523
 Patent No. 5824782
 GENERAL INFORMATION:
 APPLICANT: Hoelzer, Wolfgang
 APPLICANT: von Hoegen, Ilka
 APPLICANT: Strittmatter, Wolfgang
 APPLICANT: Matzku, Siegfried
 TITLE OF INVENTION: Immunoconjugates II
 NUMBER OF SEQUENCES: 13
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Millen, White, Zelano & Branigan, P.C.
 STREET: 2200 Clarendon Boulevard, Suite 1400
 CITY: Arlington
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22201
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/528,523
 FILING DATE: 06-NOV-1992
 CLASSIFICATION: 536

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: EP 94114572.4
;; FILING DATE: 16-SEP-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hamlet-King, Diana
;; REGISTRATION NUMBER: 33,302
;; REFERENCE/DOCKET NUMBER: Merck 1717
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 703-243-6333
;; TELEFAX: 703-243-6410
;; TELEX: 64191
;; INFORMATION FOR SEQ ID NO: 13:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 5 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; HYPOTHETICAL: NO
;; FRAGMENT TYPE: internal
;; US-08-528-523-13
Query Match 100.0%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGGS 5
Db 1 GGGGS 5
RESULT 14
US-08-537-874-62
; Sequence 62, Application US/08537874
; Patent No. 5830721
; GENERAL INFORMATION:
; APPLICANT: Stemmer, Willem P.C.
; APPLICANT: Cramer, Andreas
; TITLE OF INVENTION: DNA Mutagenesis by Random Fragmentation
; TITLE OF INVENTION: and Reassembly
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/537,874
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP PCT/US95/02126
; FILING DATE: 17-FEB-1995
; APPLICATION NUMBER: US 08/198,431
; FILING DATE: 17-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 018097-014610
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid

;; STRANDEDNESS: not relevant
;; TOPOLOGY: not relevant
;; US-08-537-874-62
Query Match 100.0%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGGS 5
Db 1 GGGGS 5
RESULT 15
US-08-448-418-86
; Sequence 86, Application US/08448418
; Patent No. 5837242
; GENERAL INFORMATION:
; APPLICANT: Holliger, Kaspar-Philipp
; APPLICANT: Griffiths, Andrew D
; APPLICANT: Hoogenboom, Hendricus RJM
; APPLICANT: Malmqvist, Magnus
; APPLICANT: Marks, James D
; APPLICANT: McGuinness, Brian T
; APPLICANT: Pope, Anthony R
; APPLICANT: Prospero, Terence D
; APPLICANT: Winter, Gregory P
; TITLE OF INVENTION: Multivalent and Multispecific Binding
; TITLE OF INVENTION: Proteins, Their Manufacture and Use
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall O'Toole Gerstein Murray and Borun
; STREET: 6300 Sears Tower 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,418
; FILING DATE: 14-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB93/02492
; FILING DATE: 03-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9225453.1
; FILING DATE: 04-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9300816.7
; FILING DATE: 16-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 93303614.7
; FILING DATE: 10-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9319969.3
; FILING DATE: 22-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: David W. Clough
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 28111/32651
; INFORMATION FOR SEQ ID NO: 86:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide linker
; US-08-448-418-86

Query Match 100.0%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2y 1 GGGGS 5
|||||
db 1 GGGGS 5

RESULT 16
JS-08-471-939-140
Sequence 140, Application US/08471939
Patent No. 5844076
GENERAL INFORMATION:
APPLICANT: Kay, B. K.
APPLICANT: Fowlkes, D. M.
TITLE OF INVENTION: Totally Synthetic Affinity Reagents
NUMBER OF SEQUENCES: 141
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,939
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/013,416
FILING DATE: 01-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION/DOCKET NUMBER: 1101-143
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TOPLOGY: unknown
MOLECULE TYPE: peptide
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MEDIUM TYPE: peptide
JS-08-471-939-140

Query Match 100.0%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 1 GGGGS 5
|||||
b 1 GGGGS 5

RESULT 17
S-08-471-800-140
Sequence 140, Application US/08471800
Patent No. 5852167
GENERAL INFORMATION:
APPLICANT: Kay, B. K.
APPLICANT: Fowlkes, D. M.
TITLE OF INVENTION: Totally Synthetic Affinity Reagents
NUMBER OF SEQUENCES: 141
CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,800
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/013,416
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION/DOCKET NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-143
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 140:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-471-800-140

Query Match 100.0%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GGGGS 5
|||||
Db 1 GGGGS 5

RESULT 18
US-08-932-589-58
Sequence 58, Application US/08932589
Patent No. 5874407
GENERAL INFORMATION:
APPLICANT: Kelley, Robert F.
APPLICANT: Lazarus, Robert A.
APPLICANT: Lee, Geoffrey F.
TITLE OF INVENTION: No. 5874407el Factor VIIa Inhibitors
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 MB floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/932,589
FILING DATE: 17-Sep-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/566800
FILING DATE: 12/04/95

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Kubinec, Jeffrey S.
; REGISTRATION NUMBER: 36,575
; REFERENCE/DOCKET NUMBER: P0958BD1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8228
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-932-589-58

Query Match 100.0%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 1 GGGGS 5

RESULT 19
US-08-647-449-26
; Sequence 26, Application US/08647449
; Patent No. 5916794
; GENERAL INFORMATION:
; APPLICANT: Chandrasegaran, Srinivasan
; TITLE OF INVENTION: METHODS FOR INACTIVATING TARGET DNA AND
; FOR DETECTING CONFORMATIONAL CHANGE IN A NUCLEIC ACID
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W., NINTH FLOOR, EAST
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/647,449
; FILING DATE: 07-MAY-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kokulis, Paul N
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: PNK/4130/216443/ASH
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-647-449-26

Query Match 100.0%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 1 GGGGS 5
```

```
RESULT 20
US-08-821-840-1
; Sequence 1, Application US/08821840
; Patent No. 5919456
; GENERAL INFORMATION:
; APPLICANT: Puri, Raj K.
; APPLICANT: Debinski, Waldemar
; APPLICANT: Pastan, Ira
; APPLICANT: Osiri, Nicholas
; TITLE OF INVENTION: IL-13 Receptor Specific Chimeric
; PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/821,840
; FILING DATE: 21-MAR-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/404,685
; FILING DATE: 15-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 015280-217100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-821-840-1

Query Match 100.0%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 1 GGGGS 5

RESULT 21
US-08-902-623-75
; Sequence 75, Application US/08902623
; Patent No. 5922545
; GENERAL INFORMATION:
; APPLICANT: MATTHEAKIS, LARRY C.
; APPLICANT: DOWER, WILLIAM J.
; TITLE OF INVENTION: IN VITRO PEPTIDE AND ANTIBODY DISPLAY
; LIBRARIES
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND AND TOWNSEND AND CREW LLP
; STREET: TWO EMBARCADERO CENTER, 8TH FLOOR
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
```

COUNTRY: U.S.A.
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/902,623
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/586,176
FILING DATE: 17-JAN-1996
APPLICATION NUMBER: US 08/300,262
FILING DATE: 02-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/144,775
FILING DATE: US 29-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/12206
FILING DATE: US 25-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: DUNN, TRACY J.
REGISTRATION NUMBER: 34,587
REFERENCE/DOCKET NUMBER: 16528X-003230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 75:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
S-08-902-623-75

Query Match 100.0%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 1 GGGGS 5
b 1 GGGGS 5

RESULT 22
S-08-471-068-140
Sequence 140, Application US/08471068
Patent No. 5948635
GENERAL INFORMATION:
APPLICANT: Kay, B. K.
APPLICANT: Fowles, D. M.
TITLE OF INVENTION: Totally Synthetic Affinity Reagents
NUMBER OF SEQUENCES: 186
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,068
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/189,331
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Mirock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-155
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 140:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-471-068-140

Query Match 100.0%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GGGGS 5
Db 1 GGGGS 5

RESULT 23
US-08-308-494A-1
Sequence 1, Application US/08308494A
Patent No. 5959083
GENERAL INFORMATION:
APPLICANT: Bosslet, Klaus
APPLICANT: Seeman, Gerhard
TITLE OF INVENTION: Tetraivalent Bispecific Receptors, The
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/308,494A
FILING DATE: 21-SEP-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/891,739
FILING DATE: 01-JUN-1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P41:8120.4
FILING DATE: 03-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kulik, David J.
REGISTRATION NUMBER: 36,576
REFERENCE/DOCKET NUMBER: 05552-1186-02000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid

```
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-308-494A-1

Query Match 100.0%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
Db 1 GGGGS 5

RESULT 24
US-08-665-202-25
; Sequence 25, Application US/08665202
; Patent No. 5977322
; GENERAL INFORMATION:
; APPLICANT: Marks, James D.
; TITLE OF INVENTION: No. 5977322el High Affinity Human Antibodies to
; TITLE OF INVENTION: Tumor Antigens
; NUMBER OF SEQUENCES: 141
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 13-JUN-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,238
; FILING DATE: 14-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,250
; FILING DATE: 15-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 02307E-061410
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-665-202-25

Query Match 100.0%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
Db 1 GGGGS 5

RESULT 25
US-08-809-668-12
; Sequence 12, Application US/08809668
; Patent No. 5980895
; GENERAL INFORMATION:
; APPLICANT: Paetan, Ira
; TITLE OF INVENTION: Immunotoxin Containing a
; TITLE OF INVENTION: Disulfide-Stabilized Antibody Fragment Joined to a
; TITLE OF INVENTION: Pseudomonas Exotoxin that Does No. 5980895 Require Proteoly
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 21-AUG-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/005,388
; FILING DATE: 13-OCT-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US96/16327
; FILING DATE: 11-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 015280-253100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-809-668-12

Query Match 100.0%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
Db 1 GGGGS 5

RESULT 26
US-08-792-553-15
; Sequence 15, Application US/08792553
; Patent No. 5981200
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Heim, Roger
; TITLE OF INVENTION: Tandem Fluorescent Protein Constructs
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & RICHARDSON P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/792,553
FILING DATE: 31-JAN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lisa A. Haile, Ph.D..
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07257/041001/UC 96-160-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-678-5070
TELEFAX: 619-678-5099
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
S-08-792-553-15

Query Match 100.0%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 1 GGGGS 5
|
b 1 GGGGS 5

RESULT 27
S-08-751-767A-24
Sequence 24, Application US/08751767A
Patent No. 5994104
GENERAL INFORMATION:
APPLICANT: ANDERSON, ROBERT J.
APPLICANT: GRANT, HUGH
APPLICANT: MACDONALD, IAN D.
TITLE OF INVENTION: INTERLUKIN-12 FUSION PROTEIN
NUMBER OF SEQUENCES: 80
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHVE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/751,767A
FILING DATE: 08-NOV-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 117-221
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164091
TELEFAX: 7038164100
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide

FEATURE2:
NAME/KEY: Peptide
LOCATION: 1..5
OTHER INFORMATION: /note= "THIS SEQUENCE IS REPEATED
FROM 2 TO 10 TIMES, PREFERABLY FROM 2 TO 5 TIMES, MORE
OTHER INFORMATION: PREFERABLY 3 OR 4 TIMES, MOST PREFERABLY 3 TIMES."
US-08-751-767A-24

Query Match 100.0%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
|
Db 1 GGGGS 5

RESULT 28
US-08-818-253-40
Sequence 40, Application US/08818253
Patent No. 5988204
GENERAL INFORMATION:
APPLICANT: Tsien, Roger Y.
APPLICANT: Miyawaki, Atsushi
TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
DETECTION OF ANALYTES
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: Fast-SEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/818,253
FILING DATE: 14-MAR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Ph.D., Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07257/043001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
OTHER INFORMATION: linker moiety; sequence repeated indefinitely
US-08-818-253-40

Query Match 100.0%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
|
Db 1 GGGGS 5

RESULT 29
US-08-722-258-54

; Sequence 54, Application US/08722258
; Patent No. 6011002
; GENERAL INFORMATION:
; APPLICANT: Pastan, Ira
; APPLICANT: Kreitman, Robert J.
; APPLICANT: Puri, Raj K.
; TITLE OF INVENTION: Circularly Permuted Ligands and
; TITLE OF INVENTION: Circularly Permuted Chimeric Molecules
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 08-JAN-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; FILING DATE: 06-APR-1995
; APPLICATION NUMBER: WO PCT/US95/04468
; FILING DATE: 06-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/225,224
; FILING DATE: 08-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 015280-193100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-722-258-54

Query Match 100.0%; Score 5; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 1 GGGGS 5

RESULT 30
US-09-397-951-12
; Sequence 12, Application US/09397951
; Patent No. 6074644
; GENERAL INFORMATION:
; APPLICANT: Pastan, Ira
; APPLICANT: Kuan, Chien-Tsun
; TITLE OF INVENTION: Immunotoxin Containing a
; TITLE OF INVENTION: Disulfide-Stabilized Antibody Fragment Joined to a
; TITLE OF INVENTION: Pseudomonas Exotoxin that Does No. 6074644 Require Proteolytic
; TITLE OF INVENTION: Activation
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California

; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 08/09/397,951
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/809,668
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US96/16327
; FILING DATE: 11-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 015280-253100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-397-951-12

Query Match 100.0%; Score 5; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 1 GGGGS 5

RESULT 31
US-08-776-271-3
; Sequence 3, Application US/08776271
; Patent No. 6083502
; GENERAL INFORMATION:
; APPLICANT: Pastan, Ira
; APPLICANT: Chang, Kai
; TITLE OF INVENTION: Mesothelin, a Differentiation Antigen
; TITLE OF INVENTION: Present on Mesothelium, Mesotheliomas and Ovarian Cancers
; TITLE OF INVENTION: and Methods and Kits for Targeting the Antigen
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,271
; FILING DATE: 01-DEC-1998
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/00224
; FILING DATE: 03-JAN-1997
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/010,166
FILING DATE: 05-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Faris, Susan K.
REGISTRATION NUMBER: 41,739
REFERENCE/DOCKET NUMBER: 015280-259100US
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
JS-08-776-271-3

Query Match 100.0%; Score 5; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 1 GGGGS 5

RESULT 32
JS-08-621-859-64
Sequence 64, Application US/08621859
Patent No. 6117679
GENERAL INFORMATION:
APPLICANT: Stemmer, Willem P.C.
APPLICANT: Cramer, Andreas M.
TITLE OF INVENTION: Methods for Generating Polynucleotides
TITLE OF INVENTION: Having Desired Characteristics by Iterative Selection and
TITLE OF INVENTION: Recombination
NUMBER OF SEQUENCES: 70
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/621,859
FILING DATE: 25-MAR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/564,955
FILING DATE: 30-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/537,874
FILING DATE: 30-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US95/02126
FILING DATE: 17-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/198,431
FILING DATE: 17-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Liebeschuetz, Joe
REGISTRATION NUMBER: 37,505
REFERENCE/DOCKET NUMBER: 018097-014612US
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-621-859-64

Query Match 100.0%; Score 5; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 1 GGGGS 5

RESULT 33
US-09-100-856A-62
Sequence 62, Application US/09100856A
Patent No. 6132970
GENERAL INFORMATION:
APPLICANT: Stemmer, Willem P.C.
APPLICANT: Cramer, Andreas
TITLE OF INVENTION: DNA Mutagenesis by Random Fragmentation
TITLE OF INVENTION: and Reassembly
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,856A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/537,874
FILING DATE:
APPLICATION NUMBER: US 08/198,431
FILING DATE: 17-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Liebeschuetz, Joe
REGISTRATION NUMBER: 37,505
REFERENCE/DOCKET NUMBER: 018097-014610
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: No. 6132970 Relevant
TOPOLOGY: No. 6132970 Relevant
US-09-100-856A-62

Query Match 100.0%; Score 5; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 1 GGGGS 5

RESULT 34

US-09-046-992-5
; Sequence 5, Application US/09046992
; Patent No. 6140066
; GENERAL INFORMATION:
; APPLICANT: Lorberboun-Galski, Haya
; APPLICANT: Yarkoni, Shai
; APPLICANT: Ben-Yehudah, Ami
; TITLE OF INVENTION: METHODS OF CANCER DIAGNOSIS
; TITLE OF INVENTION: USING A CHIMERIC TOXIN
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/046,992
; FILING DATE: 24-MAR-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 9457-0013-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-046-992-5

Query Match 100.0%; Score 5; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|||
Db 1 GGGGS 5

RESULT 35

US-09-215-035-3
; Sequence 3, Application US/09215035
; Patent No. 6153430
; GENERAL INFORMATION:
; APPLICANT: Pastan, Ira
; APPLICANT: Chang, Kai
; TITLE OF INVENTION: Mesothelin, a Differentiation Antigen
; TITLE OF INVENTION: Present on Mesothelium, Mesotheliomas and Ovarian Cancers
; TITLE OF INVENTION: and Methods and Kits for Targeting the Antigen
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/215,035
; FILING DATE: No. 6153430 yet assigned
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/776,271
; FILING DATE: 01-DEC-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/00224
; FILING DATE: 03-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/010,166
; FILING DATE: 05-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Faris, Susan X.
; REGISTRATION NUMBER: 41,739
; REFERENCE/DOCKET NUMBER: 015280-2591100S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-215-035-3

Query Match 100.0%; Score 5; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|||
Db 1 GGGGS 5

RESULT 36

US-09-075-511-64
; Sequence 64, Application US/09075511
; Patent No. 6165793
; GENERAL INFORMATION:
; APPLICANT: Stemmer, Willem P.C.
; APPLICANT: Cramer, Andreas M.
; TITLE OF INVENTION: Methods for Generating Polynucleotides
; TITLE OF INVENTION: Having Desired Characteristics by Iterative Selection and
; TITLE OF INVENTION: Recombination
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/075,511
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/621,859
; FILING DATE: 25-MAR-1996

APPLICATION NUMBER: US 08/564,955
FILING DATE: 30-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/537,874
FILING DATE: 30-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US95/02126
FILING DATE: 17-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/198,431
FILING DATE: 17-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Liebeschuetz, Joe
REGISTRATION NUMBER: 37,505
REFERENCE/DOCKET NUMBER: 018097-014612US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
S-09-075-511-64

Query Match 100.0%; Score 5; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 1 GGGGS 5
|||||
b 1 GGGGS 5

RESULT 37
S-09-099-015-64
Sequence 64, Application US/09099015
Patent No. 6180406
GENERAL INFORMATION:
APPLICANT: Stemmer, Willem P.C.
APPLICANT: Cramer, Andreas M.
TITLE OF INVENTION: Methods for Generating Polynucleotides
TITLE OF INVENTION: Having Desired Characteristics by Iterative Selection and
TITLE OF INVENTION: Recombination
NUMBER OF SEQUENCES: 70
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/099,015
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/621,859
FILING DATE: 25-MAR-1996
APPLICATION NUMBER: US 08/564,955
FILING DATE: 30-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/537,874
FILING DATE: 30-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US95/02126

FILING DATE: 17-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/198,431
FILING DATE: 17-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Liebeschuetz, Joe
REGISTRATION NUMBER: 37,505
REFERENCE/DOCKET NUMBER: 018097-014612US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-099-015-64

Query Match 100.0%; Score 5; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|||||
Db 1 GGGGS 5

RESULT 38
US-08-818-252-40
Sequence 40, Application US/08818252B
Patent No. 6197928
GENERAL INFORMATION:
APPLICANT: Tsien, Roger Y.
APPLICANT: Miyawaki, Atsushi
TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
TITLE OF INVENTION: DETECTION OF ANALYTES
FILE REFERENCE: 07257/042001
CURRENT APPLICATION NUMBER: US/08/818,252B
CURRENT FILING DATE: 1997-03-14
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 40
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Linker moiety; sequence repeated indefinitely
US-08-818-252-40

Query Match 100.0%; Score 5; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|||||
Db 1 GGGGS 5

RESULT 39
US-09-091-814-1
Sequence 1, Application US/09091814
Patent No. 6218513
GENERAL INFORMATION:
APPLICANT: Anthony-Cahill, Spencer J.
APPLICANT: Epp, Janet K
APPLICANT: Kerwin, Bruce A.
APPLICANT: Olin O., Peter
APPLICANT: Mathews J., Antony
TITLE OF INVENTION: GLOBINS CONTAINING BINDING DOMAINS
FILE REFERENCE: BXTB2005
CURRENT APPLICATION NUMBER: US/09/091,814

; CURRENT FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Amino Acid
; OTHER INFORMATION: Linker
US-09-091-814-1

Query Match 100.0%; Score 5; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|||||
Db 1 GGGGS 5

RESULT 40
US-09-281-792B-26
; Sequence 26, Application US/09281792B
; Patent No. 6285196
; GENERAL INFORMATION:
; APPLICANT: Chandrasegaran, Srinivasan
; TITLE OF INVENTION: METHODS FOR INACTIVATION TARGET DNA AND
; FOR DETECTING CONFORMATIONAL CHANGE IN A NUCLEIC ACID
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: VENABLE, BARTUER, HOWARD & CIVILETTI, LLP
; STREET: 1201 New York Avenue, N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/09/281,792B
FILING DATE: 31-Mar-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Hobbs, Ann S
REGISTRATION NUMBER: 36,930
REFERENCE/DOCKET NUMBER: 36595160594
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4800
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 26:

US-09-281-792B-26
Query Match 100.0%; Score 5; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|||||
Db 1 GGGGS 5

RESULT 41

US-09-232-863-62
; Sequence 62, Application US/09232863
; Patent No. 6277638
; GENERAL INFORMATION:
; APPLICANT: Stemmer, Willem P.C.
; APPLICANT: Cramer, Andreas
; TITLE OF INVENTION: DNA Mutagenesis by Random Fragmentation
; TITLE OF INVENTION: and Reassembly
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/09/232,863
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/537,874
FILING DATE:
APPLICATION NUMBER: EP PCT/US95/02126
FILING DATE: 17-FEB-1995
APPLICATION NUMBER: US 08/198,431
FILING DATE: 17-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Liebeschuetz, Joe
REGISTRATION NUMBER: 37,505
REFERENCE/DOCKET NUMBER: 018097-014610
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: No. 6277638 Relevant
TOPOLOGY: No. 6277638 Relevant

US-09-232-863-62
Query Match 100.0%; Score 5; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|||||
Db 1 GGGGS 5

RESULT 42
US-09-362-805-1
; Sequence 1, Application US/09362805
; Patent No. 6284223
; GENERAL INFORMATION:
; APPLICANT: Luiken, George
; TITLE OF INVENTION: METHOD FOR VIEWING DISEASED TISSUE
; TITLE OF INVENTION: LOCATED WITHIN A BODY CAVITY
; FILE REFERENCE: FLUOR1120
; CURRENT APPLICATION NUMBER: US/09/362,805
; CURRENT FILING DATE: 1999-07-28
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence

```
FEATURE:
OTHER INFORMATION: a peptide linker moiety
S-09-362-805-1

Query Match      100.0%; Score 5; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 1 GGGGS 5
b 1 GGGGS 5

RESULT 43
S-09-362-805-2
Sequence 2, Application US/09362805
Patent No. 6284223
GENERAL INFORMATION:
APPLICANT: Luiken, George
TITLE OF INVENTION: METHOD FOR VIEWING DISEASED TISSUE
TITLE OF INVENTION: LOCATED WITHIN A BODY CAVITY
FILE REFERENCE: FLOOR1120
CURRENT APPLICATION NUMBER: US/09/362.805
CURRENT FILING DATE: 1999-07-28
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: a peptide linker moiety
S-09-362-805-2

Query Match      100.0%; Score 5; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 1 GGGGS 5
b 1 GGGGS 5

RESULT 44
S-09-367-953B-26
Sequence 26, Application US/09367953B
Patent No. 6287572
GENERAL INFORMATION:
APPLICANT: Kingsman, Alan J
TITLE OF INVENTION: Anti-HIV peptides and proteins
FILE REFERENCE: 550-154
CURRENT APPLICATION NUMBER: US/09/367.953B
CURRENT FILING DATE: 1999-08-24
PRIOR APPLICATION NUMBER: PCT/GB98/00563
PRIOR FILING DATE: 1998-02-23
PRIOR APPLICATION NUMBER: GB 9703802.0
PRIOR FILING DATE: 1997-02-24
NUMBER OF SEQ ID NOS: 131
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 26
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Linker
S-09-367-953B-26

Query Match      100.0%; Score 5; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 1 GGGGS 5

FEATURE:
OTHER INFORMATION: a peptide linker moiety
S-09-362-805-1

Query Match      100.0%; Score 5; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 1 GGGGS 5
b 1 GGGGS 5

RESULT 45
US-09-133-508A-62
Sequence 62, Application US/09133508A
Patent No. 6287861
GENERAL INFORMATION:
APPLICANT: Stemmer, Willem P.C.
APPLICANT: Crameri, Andreas
TITLE OF INVENTION: DNA Mutagenesis by Random Fragmentation
TITLE OF INVENTION: and Reassembly
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/133,508A
FILING DATE: 12-AUG-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/537,874
FILING DATE: 04-MAR-1996
APPLICATION NUMBER: EP PCT/US95/02126
FILING DATE: 17-FEB-1995
APPLICATION NUMBER: US 08/198,431
FILING DATE: 17-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Liebeschuetz, Joe
REGISTRATION NUMBER: 37,505
REFERENCE/DOCKET NUMBER: 018097-014610
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: No. 6287861 Relevant
TOPOLOGY: No. 6287861 Relevant
US-09-133-508A-62

Query Match      100.0%; Score 5; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 1 GGGGS 5
b 1 GGGGS 5

RESULT 46
US-09-165-060-64
Sequence 64, Application US/09165060
Patent No. 6291242
GENERAL INFORMATION:
APPLICANT: Stemmer, Willem P.C.
APPLICANT: Crameri, Andreas M.
TITLE OF INVENTION: Methods for Generating Polynucleotides
Having Desired Characteristics by Iterative Selection and
Recombination
NUMBER OF SEQUENCES: 70
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
```

```

; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
;
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/165,060
; FILING DATE: 02-Oct-1998
; CLASSIFICATION: <Unknown>
; 30-OCT-1995
; 17-FEB-1995
; 17-FEB-1994
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/621,859
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/537,874
; FILING DATE: 30-OCT-1995
; APPLICATION NUMBER: WO PCT/US95/02126
; FILING DATE: 17-FEB-1995
; APPLICATION NUMBER: US 08/198,431
; FILING DATE: 17-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 018097-014612US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; 30-OCT-1995
; 17-FEB-1995
; 17-FEB-1994
;
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 64:
;
; US-09-165-060-64

```

```

Query Match 100.0%; Score 5; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GGGGS 5
Db 1 GGGGS 5

```

RESULT 47

```

US-09-501-698-62
; Sequence 62, Application US/09501698
; Patent No. 6297053
; GENERAL INFORMATION:
; APPLICANT: Stemmer, Willem P.C.
; APPLICANT: Cramer, Andreas
; TITLE OF INVENTION: DNA Mutagenesis by Random Fragmentation
; TITLE OF INVENTION: and Reassembly
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/501,698
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/133,508
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 018097-014610
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. 6297053 Relevant
; TOPOLOGY: No. 6297053 Relevant
; US-09-501-698-62

```

```

Query Match 100.0%; Score 5; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GGGGS 5
Db 1 GGGGS 5

```

RESULT 48

```

US-09-173-190-1
; Sequence 1, Application US/09173190
; Patent No. 6299860
; GENERAL INFORMATION:
; APPLICANT: Luiken, George
; TITLE OF INVENTION: METHOD FOR VIEWING DISEASED TISSUE
; TITLE OF INVENTION: LOCATED WITHIN A BODY CAVITY
; FILE REFERENCE: FLUOR1120
; CURRENT APPLICATION NUMBER: US/09/173,190
; CURRENT FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: a peptide linker moiety
; US-09-173-190-1

```

```

Query Match 100.0%; Score 5; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GGGGS 5
Db 1 GGGGS 5

```

RESULT 49

```

US-09-173-190-2
; Sequence 2, Application US/09173190
; Patent No. 6299860
; GENERAL INFORMATION:
; APPLICANT: Luiken, George
; TITLE OF INVENTION: METHOD FOR VIEWING DISEASED TISSUE
; TITLE OF INVENTION: LOCATED WITHIN A BODY CAVITY

```

```
FILE REFERENCE: FLUOR1120
CURRENT APPLICATION NUMBER: US/09/173,190
CURRENT FILING DATE: 1998-10-15
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 2
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: a peptide linker moiety
JS-09-173-190-2

Query Match      100.0%; Score 5; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y      1 GGGGS 5
      |||||
b      1 GGGGS 5

RESULT 50
S-09-240-310-64
Sequence 64, Application US/09240310
Patent No. 6323030
GENERAL INFORMATION:
APPLICANT: Stemmer, Willem P.C.
Crameri, Andreas M.
TITLE OF INVENTION: Methods for Generating Polynucleotides
Having Desired Characteristics by Iterative Selection and
Recombination
NUMBER OF SEQUENCES: 70
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/240,310
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/621,859
FILING DATE:
APPLICATION NUMBER: US 08/537,874
FILING DATE: 30-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US95/02126
FILING DATE: 17-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/198,431
FILING DATE: 17-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Liebeschuetz, Joe
REGISTRATION NUMBER: 37,505
REFERENCE/DOCKET NUMBER: 018097-014612US
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 64:
US-09-590-778-64

Query Match      100.0%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGGGS 5
      |||||
DB      1 GGGGS 5

RESULT 51
US-09-590-778-64
Sequence 64, Application US/09590778
Patent No. 6344356
GENERAL INFORMATION:
APPLICANT: Stemmer, Willem P.C.
Crameri, Andreas M.
TITLE OF INVENTION: Methods for Generating Polynucleotides
Having Desired Characteristics by Iterative Selection and
Recombination
NUMBER OF SEQUENCES: 70
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/590,778
FILING DATE: 08-Jun-2000
CLASSIFICATION: <Unknown>
30-OCT-1995
17-FEB-1995
17-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/621,859
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/537,874
FILING DATE: 30-OCT-1995
APPLICATION NUMBER: WO PCT/US95/02126
FILING DATE: 17-FEB-1995
APPLICATION NUMBER: US 08/198,431
FILING DATE: 17-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Liebeschuetz, Joe
REGISTRATION NUMBER: 37,505
REFERENCE/DOCKET NUMBER: 018097-014612US
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
30-OCT-1995
17-FEB-1995
17-FEB-1994
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 64:
US-09-590-778-64

Query Match      100.0%; Score 5; DB 4; Length 5;
```

Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
Db 1 GGGGS 5

RESULT 52
US-09-150-864A-18
; Sequence 18, Application US/09150864A
; Patent No. 6355779
; GENERAL INFORMATION:
; APPLICANT: Alderson, Mark R.
; APPLICANT: Goodwin, Raymond G.
; TITLE OF INVENTION: Cytokine Designated 4-1BB Ligand and Human Receptor
; FILE REFERENCE: 2801-B
; CURRENT APPLICATION NUMBER: US/09/150,864A
; CURRENT FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 08/060,843
; PRIOR FILING DATE: 1993-05-07
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: peptide linker
US-09-150-864A-18

Query Match 100.0%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
Db 1 GGGGS 5

RESULT 53
US-09-535-754-3
; Sequence 3, Application US/09535754
; Patent No. 6361974
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: DJAVAKHSHVILI, Tsotne
; TITLE OF INVENTION: EXONUCLEASE-MEDIATED NUCLEIC ACID REASSEMBLY IN DIRECTED EVOLUTIO
; FILE REFERENCE: DIVER1460-14
; CURRENT APPLICATION NUMBER: US/09/535,754
; CURRENT FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 09/522,289
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Antibody spacer peptide. The entire peptide sequence can be repe
; OTHER INFORMATION: ated more than one time
US-09-535-754-3

Query Match 100.0%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5

Db 1 GGGGS 5

RESULT 54
US-09-477-763-62
; Sequence 62, Application US/09477763
; Patent No. 6365408
; GENERAL INFORMATION:
; APPLICANT: Stemmer, Willem P.C.
; APPLICANT: Cramer, Andreas
; TITLE OF INVENTION: DNA Mutagenesis by Random Fragmentation
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/09/477,763
; FILING DATE: 04-Jan-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,856
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/198,431
; FILING DATE: 17-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 018097-014610
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. 6365408 Relevant
; TOPOLOGY: No. 6365408 Relevant
; SEQUENCE DESCRIPTION: SEQ ID NO: 62:
US-09-477-763-62

Query Match 100.0%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
Db 1 GGGGS 5

RESULT 55
US-09-215-212-1
; Sequence 1, Application US/09215212
; Patent No. 6372207
; GENERAL INFORMATION:
; APPLICANT: TEPPER, Mark
; APPLICANT: CUNNINGHAM, Mark
; APPLICANT: SHERRIS, David
; APPLICANT: EL TAYAR, Nabil
; APPLICANT: MCKENNA, Sean
; TITLE OF INVENTION: IFNAR2/IFN COMPLEX
; FILE REFERENCE: TEPPER1A.SEQ
; CURRENT APPLICATION NUMBER: US/09/215,212

```

CURRENT FILING DATE: 1998-12-18
PRIOR APPLICATION NUMBER: 60/068,295
PRIOR FILING DATE: 1997-12-19
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
TYPE: PRT
LENGTH: 5
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Linker
JS-09-215-212-1

Query Match      100.0%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 1 GGGGS 5
b 1 GGGGS 5

RESULT 56
S-09-590-774-64
Sequence 64, Application US/09590774
Patent No. 6372497
GENERAL INFORMATION:
APPLICANT: Stemmer, Willem P.C.
APPLICANT: Cramer, Andreas M.
TITLE OF INVENTION: Methods for Generating Polynucleotides
Having Desired Characteristics by Iterative Selection and
Recombination
NUMBER OF SEQUENCES: 70
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/590,774
FILING DATE: 17-FEB-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/621,859
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/537,874
FILING DATE: 30-OCT-1995
APPLICATION NUMBER: WO PCT/US95/02126
FILING DATE: 17-FEB-1994
APPLICATION NUMBER: US 08/198,431
FILING DATE: 17-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Liebeschuetz, Joe
REGISTRATION NUMBER: 37,505
REFERENCE/DOCKET NUMBER: 018097-014612US
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS:
MOLECULE TYPE: linear
TOPLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 64:
US-09-619-550-64

Query Match      100.0%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 1 GGGGS 5
b 1 GGGGS 5

TOPLOGY: linear
MOLECULE TYPE: peptide
US-09-590-774-64

Query Match      100.0%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 1 GGGGS 5
b 1 GGGGS 5

RESULT 57
US-09-619-550-64
Sequence 64, Application US/09619550
Patent No. 6395547
GENERAL INFORMATION:
APPLICANT: Stemmer, Willem P.C.
APPLICANT: Cramer, Andreas M.
TITLE OF INVENTION: Methods for Generating Polynucleotides
Having Desired Characteristics by Iterative Selection and
Recombination
NUMBER OF SEQUENCES: 70
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/619,550
FILING DATE: 19-Jul-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/621,859
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/537,874
FILING DATE: 30-OCT-1995
APPLICATION NUMBER: WO PCT/US95/02126
FILING DATE: 17-FEB-1995
APPLICATION NUMBER: US 08/198,431
FILING DATE: 17-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Liebeschuetz, Joe
REGISTRATION NUMBER: 37,505
REFERENCE/DOCKET NUMBER: 018097-014612US
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS:
MOLECULE TYPE: linear
TOPLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 64:
US-09-619-550-64

Query Match      100.0%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 1 GGGGS 5
b 1 GGGGS 5
```

```

; Cramer, Andreas M.
; TITLE OF INVENTION: Methods for Generating Polynucleotides
; Having Desired Characteristics by Iterative Selection and
; Recombination
;
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/240,734
; FILING DATE: 29-Jan-1999
; CLASSIFICATION: <Unknown>
; 30-OCT-1995
; 17-FEB-1995
; 17-FEB-1994
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/621,859
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/537,874
; FILING DATE: 30-OCT-1995
; APPLICATION NUMBER: WO PCT/US95/02126
; FILING DATE: 17-FEB-1995
; APPLICATION NUMBER: US 08/198,431
; FILING DATE: 17-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 018097-014612US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; 30-OCT-1995
; 17-FEB-1995
; 17-FEB-1994
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 64:
;
; US-09-240-734-64
;
; Query Match 100.0%; Score 5; DB 4; Length 5;
; Best Local Similarity 100.0%; Pred. No. 2.5e+05;
; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 1 GGGGS 5
; Db 1 GGGGS 5
;
; RESULT 61
; US-09-231-253-62
; Sequence 62, Application US/09231253
; Patent No. 6420175
; GENERAL INFORMATION:
; APPLICANT: Stemmer, Willem P.C.
; TITLE OF INVENTION: DNA Mutagenesis by Random Fragmentation
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:

```

```

; US-09-240-179-52
; Sequence 52, Application US/09240179
; Patent No. 6410248
; GENERAL INFORMATION:
; APPLICANT: Greisman, Harvey A.
; APPLICANT: Pabo, Carl O.
; APPLICANT: Massachusetts Institute of Technology
; TITLE OF INVENTION: A General Strategy for Selecting High-Affinity Zinc
; Finger Proteins for Diverse DNA Target Sites
; FILE REFERENCE: 019496-000220US
; CURRENT APPLICATION NUMBER: US/09/240,179
; CURRENT FILING DATE: 1999-01-29
; EARLIER APPLICATION NUMBER: US 60/073,223
; EARLIER FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:linker
;
; US-09-240-179-52
;
; Query Match 100.0%; Score 5; DB 4; Length 5;
; Best Local Similarity 100.0%; Pred. No. 2.5e+05;
; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 1 GGGGS 5
; Db 1 GGGGS 5
;
; RESULT 59
; US-09-602-373A-76
; Sequence 76, Application US/09602373A
; Patent No. 6410271
; GENERAL INFORMATION:
; APPLICANT: Zhu, Li
; APPLICANT: Rua, Shaobing B.
; TITLE OF INVENTION: GENERATION OF HIGHLY DIVERSE LIBRARY OF EXPRESSION
; VECTORS VIA HOMOLOGOUS RECOMBINATION IN YEAST
; FILE REFERENCE: 25636-702 Seq Listing
; CURRENT APPLICATION NUMBER: US/09/602,373A
; CURRENT FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 76
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: linker peptide
;
; US-09-602-373A-76
;
; Query Match 100.0%; Score 5; DB 4; Length 5;
; Best Local Similarity 100.0%; Pred. No. 2.5e+05;
; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 1 GGGGS 5
; Db 1 GGGGS 5
;
; RESULT 60
; US-09-240-734-64
; Sequence 64, Application US/09240734
; Patent No. 6413774
; GENERAL INFORMATION:
; APPLICANT: Stemmer, Willem P.C.

```


ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION NUMBER: US/09/231,253

FILING DATE: 15-Jan-1999
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICANT: Stemmer, Willem P.C.

ADDRESS: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: CA

COUNTRY: USA

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION NUMBER: US/09/724,958

FILING DATE: 28-Feb-1994

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICANT: Liebeschuetz, Joe

ADDRESS: 37,505

STREET: 018097-014610

CITY: 018097-014610

STATE: 018097-014610

COUNTRY: 018097-014610

ZIP: 018097-014610

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION NUMBER: US/09/724,958

FILING DATE: 28-Feb-1994

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICANT: Liebeschuetz, Joe

ADDRESS: 37,505

STREET: 018097-014610

CITY: 018097-014610

STATE: 018097-014610

COUNTRY: 018097-014610

ZIP: 018097-014610

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION NUMBER: US/09/724,958

FILING DATE: 28-Feb-1994

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICANT: Liebeschuetz, Joe

ADDRESS: 37,505

STREET: 018097-014610

CITY: 018097-014610

STATE: 018097-014610

COUNTRY: 018097-014610

ZIP: 018097-014610

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION NUMBER: US/09/724,958

FILING DATE: 28-Feb-1994

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICANT: Liebeschuetz, Joe

ADDRESS: 37,505

STREET: 018097-014610

CITY: 018097-014610

STATE: 018097-014610

COUNTRY: 018097-014610

QY 1 GGGGS 5
|||||
Db 1 GGGGS 5

RESULT 63

US-09-724-958-62
; Sequence 62, Application US/09724958
; Patent No. 6444468

GENERAL INFORMATION:

APPLICANT: Stemmer, Willem P.C.

ADDRESS: Crameri, Andreas

STREET: DNA Mutagenesis by Random Fragmentation

CITY: and Reassembly

STATE: CA

COUNTRY: USA

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICANT: Liebeschuetz, Joe

ADDRESS: 37,505

STREET: 018097-014610

CITY: 018097-014610

STATE: 018097-014610

COUNTRY: 018097-014610

ZIP: 018097-014610

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICANT: Liebeschuetz, Joe

ADDRESS: 37,505

STREET: 018097-014610

CITY: 018097-014610

STATE: 018097-014610

COUNTRY: 018097-014610

ZIP: 018097-014610

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICANT: Liebeschuetz, Joe

ADDRESS: 37,505

STREET: 018097-014610

CITY: 018097-014610

STATE: 018097-014610

COUNTRY: 018097-014610

ZIP: 018097-014610

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICANT: Liebeschuetz, Joe

ADDRESS: 37,505

STREET: 018097-014610

CITY: 018097-014610

STATE: 018097-014610

COUNTRY: 018097-014610

ZIP: 018097-014610

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICANT: Liebeschuetz, Joe

ADDRESS: 37,505

STREET: 018097-014610

CITY: 018097-014610

STATE: 018097-014610

COUNTRY: 018097-014610

ZIP: 018097-014610

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

QY 1 GGGGS 5
|||||
Db 1 GGGGS 5

RESULT 64

US-09-229-007A-3
; Sequence 3, Application US/09229007A
; Patent No. 6453242

GENERAL INFORMATION:

APPLICANT: Eisenberg, Stephen P.

ADDRESS: Cox III, George N.

STREET: Jamieson, Andrew

CITY: Rebar, Edward J.

STATE: Sangamo Biosciences, Inc.

COUNTRY: Selection of Sites for Targeting by Zinc Finger

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICANT: Liebeschuetz, Joe

ADDRESS: 37,505

STREET: 018097-014610

CITY: 018097-014610

STATE: 018097-014610

COUNTRY: 018097-014610

ZIP: 018097-014610

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICANT: Liebeschuetz, Joe

ADDRESS: 37,505

STREET: 018097-014610

CITY: 018097-014610

STATE: 018097-014610

COUNTRY: 018097-014610

ZIP: 018097-014610

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICANT: Liebeschuetz, Joe

ADDRESS: 37,505

STREET: 018097-014610

CITY: 018097-014610

STATE: 018097-014610

COUNTRY: 018097-014610

ZIP: 018097-014610

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICANT: Liebeschuetz, Joe

ADDRESS: 37,505

STREET: 018097-014610

CITY: 018097-014610

STATE: 018097-014610

COUNTRY: 018097-014610

ZIP: 018097-014610

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICANT: Liebeschuetz, Joe

ADDRESS: 37,505

STREET: 018097-014610

CITY: 018097-014610

STATE: 018097-014610

COUNTRY: 018097-014610

ZIP: 018097-014610

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

QY 1 GGGGS 5
|||||
Db 1 GGGGS 5

RESULT 64

US-09-229-007A-3
; Sequence 3, Application US/09229007A
; Patent No. 6453242

GENERAL INFORMATION:

APPLICANT: Eisenberg, Stephen P.

ADDRESS: Cox III, George N.

STREET: Jamieson, Andrew

CITY: Rebar, Edward J.

STATE: Sangamo Biosciences, Inc.

COUNTRY: Selection of Sites for Targeting by Zinc Finger

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICANT: Liebeschuetz, Joe

ADDRESS: 37,505

STREET: 018097-014610

CITY: 018097-014610

STATE: 018097-014610

COUNTRY: 018097-014610

ZIP: 018097-014610

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICANT: Liebeschuetz, Joe

ADDRESS: 37,505

STREET: 018097-014610

CITY: 018097-014610

STATE: 018097-014610

COUNTRY: 018097-014610

ZIP: 018097-014610

COMPUTER READABLE FORM:

```
; TITLE OF INVENTION: Proteins and Methods of Designing Zinc Finger Proteins
; TITLE OF INVENTION: to Bind to Preslected Sites
; FILE REFERENCE: 019496-001800US
; CURRENT APPLICATION NUMBER: US/09/229,007A
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:peptide linker
US-09-229-007A-3

Query Match          100.0%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
DB 1 GGGGS 5

RESULT 65
US-09-300-008B-64
; Sequence 64, Application US/091000008B
; Patent No. 6458534
; GENERAL INFORMATION:
; APPLICANT: Concannon et al.
; TITLE OF INVENTION: A GENE ASSOCIATED WITH NIJMEGEN BREAKAGE
; TITLE OF INVENTION: SYNDROME, ITS GENE PRODUCT AND METHODS FOR THEIR USE
; FILE REFERENCE: 9924-0003-228
; CURRENT APPLICATION NUMBER: US/09/300,008B
; CURRENT FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: US 60/083,269
; PRIOR FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 64
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Flexible polylinker
US-09-300-008B-64

Query Match          100.0%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
DB 1 GGGGS 5

RESULT 66
US-09-470-191-95
; Sequence 95, Application US/09470191
; Patent No. 6465633
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods of Their Use in
; TITLE OF INVENTION: the Treatment, Prevention and diagnosis of Tuberculosis
; FILE REFERENCE: 014058-008910US
; CURRENT APPLICATION NUMBER: US/09/470,191
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 60/113,952
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 95
```

```
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: flexible polylinker
US-09-470-191-95

Query Match          100.0%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
DB 1 GGGGS 5

RESULT 67
US-09-316-919-56
; Sequence 56, Application US/09116919
; Patent No. 6469154
; GENERAL INFORMATION:
; APPLICANT: Taten, Roger Y.
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: FLUORESCENT PROTEIN INDICATORS
; FILE REFERENCE: 07257/073001
; CURRENT APPLICATION NUMBER: US/09/316,919
; CURRENT FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Linker moiety; sequence repeated indefinitely
US-09-316-919-56

Query Match          100.0%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
DB 1 GGGGS 5

RESULT 68
US-09-146-979-86
; Sequence 86, Application US/09146979
; Patent No. 6492123
; GENERAL INFORMATION:
; APPLICANT: Holliger, Kaspar-Philipp
; APPLICANT: Griffiths, Andrew D
; APPLICANT: Hoogenboom, Hendricus RJM
; APPLICANT: Malmqvist, Magnus
; APPLICANT: Marks, James D
; APPLICANT: McGuinness, Brian T
; APPLICANT: Pope, Anthony R
; APPLICANT: Prospero, Terence D
; APPLICANT: Winter, Gregory P
; TITLE OF INVENTION: Multivalent and Multispecific Binding
; TITLE OF INVENTION: Proteins, Their Manufacture and Use
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall O'Toole Gerstein Murray and Borun
; STREET: 6300 Sears Tower 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/146,979
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/448,418
FILING DATE: 14-MAY-1996
APPLICATION NUMBER: PCT/GB93/02492
FILING DATE: 03-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9225453.1
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9300816.7
FILING DATE: 16-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 93303614.7
FILING DATE: 10-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9319969.3
FILING DATE: 22-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: David W. Clough
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 28111/32651
INFORMATION FOR SEQ ID NO: 86:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide linker
JS-09-146-979-86
Query Match 100.0%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGGS 5
DB 1 GGGGS 5
RESULT 69
US-08-840-713-43
Sequence 43, Application US/08840713
Patent No. 6498233
GENERAL INFORMATION:
APPLICANT: WELLS, Winfried, Dr.
APPLICANT: FOYMINAYA, Jesus
TITLE OF INVENTION: NUCLEIC ACID TRANSFER SYSTEM
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
STREET: 655 15th St., N.W., Suite 330 - G St. Lobby
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/840,713
FILING DATE: 25-APR-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kitts, Monica Chin
REGISTRATION NUMBER: 36,105
REFERENCE/DOCKET NUMBER: 1614-7014

TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 638 - 5000
TELEFAX: (202) 638 - 4810
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-840-713-43
Query Match 100.0%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGGS 5
DB 1 GGGGS 5
RESULT 70
US-09-761-962A-36
Sequence 36, Application US/09761962A
Patent No. 6500927
GENERAL INFORMATION:
APPLICANT: Memorial Sloan-Kettering Cancer Center
TITLE OF INVENTION: MULTIPLE SPLICE VARIANTS OF THE MU-OPIOID RECEPTOR GENE
FILE REFERENCE: 830002-2000.2
CURRENT APPLICATION NUMBER: US/09/761,962A
CURRENT FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 09/743,872
PRIOR APPLICATION NUMBER: PCT/US99/15974
PRIOR FILING DATE: 1999-07-15
NUMBER OF SEQ ID NOS: 46
SOFTWARE: PatentIn version 3.0
SEQ ID NO 36
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: basic unit of a linking peptide
US-09-761-962A-36
Query Match 100.0%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGGS 5
DB 1 GGGGS 5
RESULT 71
US-09-731-558-6
Sequence 6, Application US/09731558
Patent No. 6503717
GENERAL INFORMATION:
APPLICANT: Case, Casey Christopher
APPLICANT: Liu, Qiang
APPLICANT: Rebar, Edward J.
APPLICANT: Sangamo Biosciences, Inc.
TITLE OF INVENTION: Methods of Using Randomized Libraries of Zinc Finger
FILE REFERENCE: 019496-0032100S
CURRENT APPLICATION NUMBER: US/09/731,558
CURRENT FILING DATE: 2000-12-06
PRIOR APPLICATION NUMBER: US 09/456,100
PRIOR FILING DATE: 1999-12-06
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 5
TYPE: PRT

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: polypeptide
; OTHER INFORMATION: linker
US-09-731-558-6
  Query Match      100.0%; Score 5; DB 4; Length 5;
  Best Local Similarity 100.0%; Pred. No. 2.5e+05;
  Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGGGS 5
       |||||
Db      1 GGGGS 5

RESULT 73
US-09-717-391-64
; Sequence 64, Application US/09717391
; Patent No. 6506603
; GENERAL INFORMATION:
; APPLICANT: Stemmer, Willem P.C.
; Cramer, Andreas M.
; TITLE OF INVENTION: Methods for Generating Polynucleotides
; Having Desired Characteristics by Iterative Selection and
; Recombination
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/717,391
; FILING DATE: 18-Oct-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/621,859
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/537,874
; FILING DATE: 30-Oct-1995
; APPLICATION NUMBER: WO PCT/US95/02126
; FILING DATE: 17-FEB-1995
; APPLICATION NUMBER: US 08/198,431
; FILING DATE: 17-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 018097-014612US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 64:
US-09-717-391-64

  Query Match      100.0%; Score 5; DB 4; Length 5;
  Best Local Similarity 100.0%; Pred. No. 2.5e+05;
  Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGGGS 5
       |||||
Db      1 GGGGS 5

RESULT 74
US-09-315-574-25
; Sequence 25, Application US/09315574
; Patent No. 6512097
; GENERAL INFORMATION:
; APPLICANT: Marks, James D.
; ATTORNEY/AGENT INFORMATION:
; NAME: Schier, Robert
```

TITLE OF INVENTION: No. 6512097el High Affinity Human Antibodies to
 TITLE OF INVENTION: Tumor Antigens
 NUMBER OF SEQUENCES: 141

CORRESPONDENCE ADDRESS:

ADDRESSEE: Majestic, Parsons, Siebert & Hsue P.C.

STREET: Four Embarcadero Center, Suite 1100

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-4106

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/315,574

FILING DATE: 20-MAY-99

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/000,238

FILING DATE: 14-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/000,250

FILING DATE: 15-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/665,202

FILING DATE: 13-JUN-1996

ATTORNEY/AGENT INFORMATION:

NAME: Hunter, Tom

REGISTRATION NUMBER: 38,498

REFERENCE/DOCKET NUMBER: 02307E-061411

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

S-09-315-574-25

Query Match 100.0%; Score 5; DB 4; Length 5;

Best Local Similarity 100.0%; Pred. No. 2.5e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 1 GGGGS 5

|||||

b 1 GGGGS 5

RESULT 75

S-08-913-370-1

Sequence 1, Application US/08913370

Patent No. 6518061

GENERAL INFORMATION:

APPLICANT: Puri, Raj K.

APPLICANT: Debinski, Waldemar

APPLICANT: Pastan, Ira

APPLICANT: Obiri, Nicholas

TITLE OF INVENTION: IL-13 Receptor Specific Chimeric

TITLE OF INVENTION: Proteins and Uses Thereof

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/913,370

FILING DATE: 17-FEB-1998

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/404,685

FILING DATE: 15-MAR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US96/03486

FILING DATE: 15-MAR-1996

ATTORNEY/AGENT INFORMATION:

NAME: Hunter, Tom

REGISTRATION NUMBER: 38,498

REFERENCE/DOCKET NUMBER: 015280-217110US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-913-370-1

Query Match 100.0%; Score 5; DB 4; Length 5;

Best Local Similarity 100.0%; Pred. No. 2.5e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5

|||||

Db 1 GGGGS 5

RESULT 76

US-09-696-313-64

Sequence 64, Application US/09696313

Patent No. 6518065

GENERAL INFORMATION:

APPLICANT: Stemmer, Willem P.C.

APPLICANT: Crameri, Andreas M.

TITLE OF INVENTION: Methods for Generating Polynucleotides

TITLE OF INVENTION: Having Desired Characteristics by Iterative Selection and

TITLE OF INVENTION: Recombination

NUMBER OF SEQUENCES: 70

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/696,313

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/621,859

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/537,874

FILING DATE: 30-OCT-1995

PRIOR APPLICATION DATA:

;
; APPLICATION NUMBER: WO PCT/US95/02126
; FILING DATE: 17-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/198,431
; FILING DATE: 17-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 018097-014612US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-696-313-64

Query Match 100.0%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
| | | | |
Db 1 GGGGS 5

RESULT 77
US-09-077-028A-17
; Sequence 17, Application US/09077028A
; Patent No. 6531133
; GENERAL INFORMATION:
; APPLICANT: HAYA LOBERBOOM-GALSKI
; APPLICANT: IDA STEINERGER
; APPLICANT: EVELINE BERAUD
; APPLICANT: IRINA MARIANOVSKY
; APPLICANT: SHAI YARKONI
; TITLE OF INVENTION: PSEUDOMONAS EXOTOXIN-MYELIN BASIC PROTEIN CHIMERIC
; FILE REFERENCE: LOBERBOOM=1
; CURRENT APPLICATION NUMBER: US/09/077,028A
; CURRENT FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: IL116044
; PRIOR FILING DATE: 1995-11-17
; PRIOR APPLICATION NUMBER: IL116559
; PRIOR FILING DATE: 1995-12-26
; PRIOR APPLICATION NUMBER: PC7/IL96/00151
; PRIOR FILING DATE: 1996-11-17
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Synthetic
US-09-077-028A-17

Query Match 100.0%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
| | | | |
Db 1 GGGGS 5

RESULT 78
US-09-229-037-8
; Sequence 8, Application US/09229037A
; Patent No. 6534261

;
; GENERAL INFORMATION:
; APPLICANT: Cox III, George No. 6534261bert
; APPLICANT: Case, Casey Christopher
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric Edward
; APPLICANT: Spratt, Sharon Kaye
; APPLICANT: Sangamo Biosciences, Inc.
; TITLE OF INVENTION: Regulation of Endogenous Gene Expression in Cells Using
; TITLE OF INVENTION: Zinc Finger Proteins
; FILE REFERENCE: 019496-002200US
; CURRENT APPLICATION NUMBER: US/09/229,037A
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:linker
US-09-229-037-8

Query Match 100.0%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
| | | | |
Db 1 GGGGS 5

RESULT 79
US-08-406-824A-22
; Sequence 22, Application US/08406824A
; Patent No. 6541610
; GENERAL INFORMATION:
; APPLICANT: SMITH, Craig A.
; TITLE OF INVENTION: TUMOR NECROSIS FACTOR-ALPHA AND BETA-RECEPTORS
; FILE REFERENCE: A-71592
; CURRENT APPLICATION NUMBER: US/08/406,824A
; CURRENT FILING DATE: 1995-03-20
; PRIOR APPLICATION NUMBER: US 08/255,849
; PRIOR FILING DATE: 1994-06-08
; PRIOR APPLICATION NUMBER: US 07/860,710
; PRIOR FILING DATE: 1992-03-30
; PRIOR APPLICATION NUMBER: US 07/523,635
; PRIOR FILING DATE: 1990-05-10
; PRIOR APPLICATION NUMBER: US 07/421,417
; PRIOR FILING DATE: 1989-10-13
; PRIOR APPLICATION NUMBER: US 07/405,370
; PRIOR FILING DATE: 1989-09-11
; PRIOR APPLICATION NUMBER: US 07/403,241
; PRIOR FILING DATE: 1989-09-05
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Miscellaneous Structure
US-08-406-824A-22

Query Match 100.0%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
| | | | |
Db 1 GGGGS 5

RESULT 80

```

JS-09-223-040-8
; Sequence 8, Application US/09223040
; Patent No. 6544522
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-009010US
; CURRENT APPLICATION NUMBER: US/09/223,040
; CURRENT FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: flexible
; OTHER INFORMATION: polylinker
JS-09-223-040-8

Query Match 100.0%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 1 GGGGS 5

RESULT 81
JS-09-240-307-64
; Sequence 64, Application US/09240307
; Patent No. 6573098
; GENERAL INFORMATION:
; APPLICANT: Stemmer, Willem P.C.
; APPLICANT: Crameri, Andreas M.
; TITLE OF INVENTION: Methods for Generating Polynucleotides
; TITLE OF INVENTION: Having Desired Characteristics by Iterative Selection and
; TITLE OF INVENTION: Recombination
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/240,307
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/621,859
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/537,874
; FILING DATE: 30-OCT-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US95/02126
; FILING DATE: 17-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/198,431
; FILING DATE: 17-FEB-1994
; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 018097-014612US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-240-307-64

Query Match 100.0%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 1 GGGGS 5

RESULT 82
US-09-640-1980-34
; Sequence 34, Application US/09640198D
; Patent No. 6586411
; GENERAL INFORMATION:
; APPLICANT: Russell, Stephen
; APPLICANT: Kay Whyte, Peng
; TITLE OF INVENTION: System for Monitoring the Location of
; TITLE OF INVENTION: Transgenes
; FILE REFERENCE: 07039-295001
; CURRENT APPLICATION NUMBER: US/09/640,198D
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 60/149,168
; PRIOR FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: No. 6586411-Cleavable Linker Site
; US-09-640-1980-34

Query Match 100.0%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 1 GGGGS 5

RESULT 83
PCT-US92-08634-6
; Sequence 6, Application PC/TUS9208634
; GENERAL INFORMATION:
; APPLICANT: Smith, Richard K.
; APPLICANT: Koduri, Raju
; APPLICANT: Young, Stephen G.
; APPLICANT: Witzum, Joseph L.
; APPLICANT: Curtiss, Linda K.
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &
; ADDRESSEE: Milnamow, Ltd.
; STREET: 180 North Stetson, Suite 4700

```

; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08634
; FILING DATE: 19921009
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,706
; FILING DATE: 18-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Gamson, Edward P.
; REGISTRATION NUMBER: 29,381
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)616-5400
; TELEFAX: (312)616-5460
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US92-08634-6

Query Match 100.0%; Score 5; DB 5; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 GGGGS 5
|||||
Db 1 GGGGS 5

RESULT 84
PCT-US95-04468-54
; Sequence 54, Application PC/TUS9504468
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: CIRCULARLY PERMUTATED LIGANDS AND
; TITLE OF INVENTION: CIRCULARLY PERMUTATED FUSION PROTEINS
; NUMBER OF SEQUENCES: 59
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04468
; FILING DATE: 07-APR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/225,224
; FILING DATE: 08-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen L.
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 15280-193-1PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown

; MOLECULE TYPE: peptide
; PCT-US95-04468-54
; Query Match 100.0%; Score 5; DB 5; Length 5;
; Best Local Similarity 100.0%; Pred. No. 2.5e+05;
; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
; Cy 1 GGGGS 5
; |||||
; Db 1 GGGGS 5
; RESULT 85
; US-08-704-170-76
; Sequence 76, Application US/08704170
; Patent No. 5707626
; GENERAL INFORMATION:
; APPLICANT: Douvas, Angeline
; APPLICANT: Takehana, Yoshi
; APPLICANT: Ehresmann, Glenn
; TITLE OF INVENTION: THERAPEUTIC STRATEGIES FOR
; TITLE OF INVENTION: IMMUNOINFECTION CLUSTER VIRUS INFECTIONS
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Robbins, Berliner & Carson
; STREET: 201 No. 5707626th Figueroa Street, Suite 500
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90012
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/704,170
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/029,850
; FILING DATE: 11-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Spitals, John P.
; REGISTRATION NUMBER: 29,215
; REFERENCE/DOCKET NUMBER: 1920-331
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 977-1001
; TELEFAX: (213) 977-1003
; INFORMATION FOR SEQ ID NO: 76:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-704-170-76

Query Match 100.0%; Score 5; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 GGGGS 5
|||||
Db 2 GGGGS 6

RESULT 86
US-08-809-668-10
; Sequence 10, Application US/08809668
; Patent No. 5980895
; GENERAL INFORMATION:
; APPLICANT: Pastan, Ira
; APPLICANT: Kuan, Chien-Tsun

;; TITLE OF INVENTION: Immunotoxin Containing a
;; TITLE OF INVENTION: Disulfide-Stabilized Antibody Fragment Joined to a
;; TITLE OF INVENTION: Pseudomonas Exotoxin that Does No. 5980895 Require Proteolytic
;; TITLE OF INVENTION: Activation
;; NUMBER OF SEQUENCES: 13
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Townsend and Townsend and Crew LLP
;; STREET: Two Embarcadero Center, Eighth Floor
;; CITY: San Francisco
;; STATE: California
;; COUNTRY: USA
;; ZIP: 94111-3834
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/809,668
;; FILING DATE: 21-AUG-1997
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 60/005,388
;; FILING DATE: 13-OCT-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: WO PCT/US96/16327
;; FILING DATE: 11-OCT-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hunter, Tom
;; REGISTRATION NUMBER: 38,498
;; REFERENCE/DOCKET NUMBER: 015280-2531000S
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 10:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 6 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; JS-08-809-668-10
;;
Query Match 100.0%; Score 5; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0;
2y 1 GGGGS 5
|||
3b 2 GGGGS 6
|||
RESULT 87
JS-09-397-951-10
Sequence 10, Application US/09397951
Patent No. 6074644
GENERAL INFORMATION:
;; APPLICANT: Pastan, Ira
;; APPLICANT: Kuan, Chien-Teun
;; TITLE OF INVENTION: Immunotoxin Containing a
;; TITLE OF INVENTION: Disulfide-Stabilized Antibody Fragment Joined to a
;; TITLE OF INVENTION: Pseudomonas Exotoxin that Does No. 6074644 Require Proteolytic
;; TITLE OF INVENTION: Activation
;; NUMBER OF SEQUENCES: 13
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Townsend and Townsend and Crew LLP
;; STREET: Two Embarcadero Center, Eighth Floor
;; CITY: San Francisco
;; STATE: California
;; COUNTRY: USA
;; ZIP: 94111-3834
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/397,951
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/809,668
;; FILING DATE:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: WO PCT/US96/16327
;; FILING DATE: 11-OCT-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hunter, Tom
;; REGISTRATION NUMBER: 38,498
;; REFERENCE/DOCKET NUMBER: 015280-2531000S
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 10:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 6 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-09-397-951-10
;;
Query Match 100.0%; Score 5; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0;
Qy 1 GGGGS 5
|||
Db 2 GGGGS 6
|||
RESULT 88
PCT-US94-02631-76
Sequence 76, Application PC/TUS9402631
GENERAL INFORMATION:
;; APPLICANT: Douvas, Angeline
;; APPLICANT: Takehana, Yoshi
;; APPLICANT: Ehresmann, Glenn
;; TITLE OF INVENTION: THERAPEUTIC STRATEGIES FOR
;; TITLE OF INVENTION: IMMUNOINFECTIVE CLUSTER VIRUS INFECTIONS
;; NUMBER OF SEQUENCES: 121
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Robbins, Berliner & Carson
;; STREET: 201 North Figueroa Street, Suite 500
;; CITY: Los Angeles
;; STATE: California
;; COUNTRY: U.S.A.
;; ZIP: 90012
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US94/02631
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/029,850
;; FILING DATE: 11-MAR-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Spitale, John P.
;; REGISTRATION NUMBER: 29,215
;; REFERENCE/DOCKET NUMBER: 1920-331
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (213) 977-1001

TELEFAX: (213) 977-1003
INFORMATION FOR SEQ ID NO: 76:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US94-02631-76

Query Match 100.0%; Score 5; DB 5; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGGS 5
DB 2 GGGGS 6

RESULT 89

US-08-902-623-49
Sequence 49, Application US/08902623
Patent No. 5922545
GENERAL INFORMATION:
APPLICANT: MATTHEAKIS, LARRY C.
APPLICANT: DOWER, WILLIAM J.
TITLE OF INVENTION: IN VITRO PEPTIDE AND ANTIBODY DISPLAY
TITLE OF INVENTION: LIBRARIES
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND AND TOWNSEND AND CREW LLP
STREET: TWO EMBARCADERO CENTER, 8TH FLOOR
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: U.S.A.
ZIP: 94111-3874
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/902,623
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/586,176
FILING DATE: 17-JAN-1996
APPLICATION NUMBER: US 08/300,262
FILING DATE: 02-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/144,775
FILING DATE: US 29-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/12206
FILING DATE: US 25-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: DUNN, TRACY J.
REGISTRATION NUMBER: 34,587
REFERENCE/DOCKET NUMBER: 16528X-003230JUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
US-08-902-623-49

Query Match 100.0%; Score 5; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGGS 5
DB 3 GGGGS 7

RESULT 90

US-09-374-454-16
Sequence 16, Application US/09374454
Patent No. 6395548
GENERAL INFORMATION:
APPLICANT: Lee, Mu-En
APPLICANT: Maemura, Koji
APPLICANT: Hsieh, Chung-Ming
TITLE OF INVENTION: METHODS OF MODULATING OF ANGIOGENESIS
FILE REFERENCE: 05433/037001
CURRENT APPLICATION NUMBER: US/09/374,454
CURRENT FILING DATE: 1999-08-13
EARLIER APPLICATION NUMBER: US 60/096,515
EARLIER FILING DATE: 1998-08-14
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16
LENGTH: 7
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Synthetically generated peptide
US-09-374-454-16

Query Match 100.0%; Score 5; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGGS 5
DB 1 GGGGS 5

RESULT 91

US-08-897-956A-38
Sequence 38, Application US/08897956A
Patent No. 6423512
GENERAL INFORMATION:
APPLICANT: Mary Ellen Digan
APPLICANT: Philip Lake
APPLICANT: Hermann Gram
TITLE OF INVENTION: Fusion Polypeptides
FILE REFERENCE: 600-7244/CPA
CURRENT APPLICATION NUMBER: US/08/897,956A
CURRENT FILING DATE: 1997-07-21
PRIOR APPLICATION NUMBER: 60/022,689
PRIOR FILING DATE: 1996-07-26
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 38
LENGTH: 7
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Linker
US-08-897-956A-38

Query Match 100.0%; Score 5; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGGS 5
DB 3 GGGGS 7

RESULT 92

```

US-08-750-128-6
; Sequence 6, Application US/08750128
; Patent No. 5916773
; GENERAL INFORMATION:
; APPLICANT: MELE Antonio,
; APPLICANT: DE SANTIS Rita,
; APPLICANT: CARLONI Cristina,
; APPLICANT: COSARELLA Annamaria
; TITLE OF INVENTION: Hybrid molecule of formula GM-CSF-L-EPO or
; TITLE OF INVENTION: EPO-L-GM-CSF for hematopoietic stimulation
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LOWE, PRICE, LeBLANC & BECKER
; STREET: 99 Canal Center Plaza, Suite 300
; CITY: Alexandria
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk, 3.50 inch.
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/750,128
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IT FI 94 A 000106
; FILING DATE: 27-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Robert L. Price
; REGISTRATION NUMBER: 22,685
; REFERENCE/DOCKET NUMBER: 2879-007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-684-1111
; TELEFAX: 703-684-1124
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; US-08-750-128-6

Query Match 100.0%; Score 5; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
Db 4 GGGGS 8
|||||

RESULT 93
US-08-487-890A-144
; Sequence 144, Application US/08487890A
; Patent No. 5708149
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena
; APPLICANT: Harkness, Robin
; APPLICANT: Schryvers, Anthony
; APPLICANT: Chong, Pele
; APPLICANT: Gray-Owen, Scott
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Murdin, Andrew
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Transferrin Receptor Genes
; NUMBER OF SEQUENCES: 147
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,890A

```

```

; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,890A
; FILING DATE: 07-JUN-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/175,116
; FILING DATE: 29-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,968
; FILING DATE: 08-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-466 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 144:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-487-890A-144

Query Match 100.0%; Score 5; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
Db 3 GGGGS 7
|||||

RESULT 94
US-08-487-890A-145
; Sequence 145, Application US/08487890A
; Patent No. 5708149
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena
; APPLICANT: Harkness, Robin
; APPLICANT: Schryvers, Anthony
; APPLICANT: Chong, Pele
; APPLICANT: Gray-Owen, Scott
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Murdin, Andrew
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Transferrin Receptor Genes
; NUMBER OF SEQUENCES: 147
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,890A

```

```

; FILING DATE: 07-JUN-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/175,116
; FILING DATE: 29-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,968
; FILING DATE: 08-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-466 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 145:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide linker
; US-08-487-890A-145

Query Match 100.0%; Score 5; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 3 GGGGS 7

RESULT 95
US-08-448-418-88
; Sequence 88, Application US/08448418
; Patent No. 5837242
; GENERAL INFORMATION:
; APPLICANT: Holliger, Kaspar-Philipp
; APPLICANT: Griffiths, Andrew D
; APPLICANT: Hoogenboom, Hendricus RJM
; APPLICANT: Malmqvist, Magnus
; APPLICANT: Marks, James D
; APPLICANT: McGuinness, Brian T
; APPLICANT: Pope, Anthony R
; APPLICANT: Prospero, Terence D
; APPLICANT: Winter, Gregory P
; TITLE OF INVENTION: Multivalent and Multispecific Binding
; TITLE OF INVENTION: Proteins, Their Manufacture and Use
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall O'Toole Gerstein, Murray and Borun
; STREET: 6300 Sears Tower 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,418
; FILING DATE: 14-MAY-1996
; CLASSIFICATION: 435
; CLASSIFICATION: C12N 15/62, 15/70, C07K 1/00
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB93/02492
; FILING DATE: 03-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9225453.1
; FILING DATE: 04-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9300816.7

; FILING DATE: 16-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 93303614.7
; FILING DATE: 10-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9319969.3
; FILING DATE: 22-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: David W. Clough
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 28111/32651
; INFORMATION FOR SEQ ID NO: 88:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide linker
; US-08-448-418-88

Query Match 100.0%; Score 5; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 3 GGGGS 7

RESULT 96
US-08-478-435-144
; Sequence 144, Application US/08478435
; Patent No. 592323
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena
; APPLICANT: Harkness, Robin
; APPLICANT: Schryvers, Anthony
; APPLICANT: Chong, Pele
; APPLICANT: Gray-Owen, Scott
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Murdin, Andrew
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Transferrin Receptor Genes
; NUMBER OF SEQUENCES: 147
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,435
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/337,483
; FILING DATE: 08-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/175,116
; FILING DATE: 29-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,968
; FILING DATE: 08-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-462 MIS:vg
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:11:55 ; Search time 42.7778 Seconds
(without alignments)
18.552 Million cell updates/sec

Title: US-10-057-890A-16

Perfect score: 5

Sequence: 1 GGGGS 5

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 251420

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Listing first 150 summaries

Database : A_Geneseq_19Jun03.*

```
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	5	100.0	5 14 AAR34034	Linking sequence w
2	5	100.0	5 16 AAR72707	Linker for apo A-I
3	5	100.0	5 17 AAR95062	scFv spacer peptid
4	5	100.0	5 18 AA17094	Gly(4)-Ser linker
5	5	100.0	5 18 AA17094	Chimeric protein p
6	5	100.0	5 20 AA19543	Linker for dual av
7	5	100.0	5 20 AA19543	VH-VL domain linke
8	5	100.0	5 20 AA19543	IFNAR2/IFN-beta co
9	5	100.0	5 20 AA19543	Peptide linker use

10	5	100.0	5 21 AAB00156	Linker used in scd
11	5	100.0	5 21 AAB14535	Peptide linker for
12	5	100.0	5 21 AAB06226	Expression vector
13	5	100.0	5 21 AAY83210	Peptide linker use
14	5	100.0	5 21 AAY43750	Linker used to con
15	5	100.0	5 21 AAY54917	Linker from IL-12
16	5	100.0	5 22 AAB56480	Human single chain
17	5	100.0	5 22 AAE11873	Spacer peptide use
18	5	100.0	5 22 AAE08730	Peptide linker #9
19	5	100.0	5 22 AAU05172	Pain-relieving tar
20	5	100.0	5 22 AAU05173	Pain-relieving tar
21	5	100.0	5 22 AAG63011	Peptide linker use
22	5	100.0	5 22 AAG06269	Glycine linker #2
23	5	100.0	5 22 AAG62343	Linker peptide SEQ
24	5	100.0	5 22 AAG62613	Linker peptide. S
25	5	100.0	5 22 AAB81033	Linker peptide use
26	5	100.0	5 22 AAB74580	Context-dependent
27	5	100.0	5 22 AAB68571	Linker peptide #1.
28	5	100.0	5 22 AAB50848	Fluorescent protei
29	5	100.0	5 22 AAB55527	Peptide linker #1.
30	5	100.0	5 22 AAB57374	Flexible polylinke
31	5	100.0	5 22 AAB57784	Flexible polylinke
32	5	100.0	5 23 AAB56192	Linker peptide. S
33	5	100.0	5 23 AAE28334	Linker peptide #5
34	5	100.0	5 23 AAB28173	Linker peptide use
35	5	100.0	5 23 ABB79370	Linker peptide use
36	5	100.0	5 23 ABB79371	Linker peptide use
37	5	100.0	5 23 AEG80672	Synthetic linker p
38	5	100.0	5 23 AAE25959	Linker peptide #3
39	5	100.0	5 23 AAB70506	VH and VL antibody
40	5	100.0	5 23 AAB31035	Binding domain-imm
41	5	100.0	5 23 AAB31505	Flexible peptide 1
42	5	100.0	5 23 AAB31871	Spacer peptide use
43	5	100.0	5 23 AAB52872	Flexible peptide 1
44	5	100.0	5 23 AAO22207	Linker peptide for
45	5	100.0	5 23 AAB31102	5 amino acid flexi
46	5	100.0	5 23 AAB83546	Peptide linker #2.
47	5	100.0	5 23 AAB703945	Peptide linker SEQ
48	5	100.0	5 23 AAB48193	Peptide linker SEQ
49	5	100.0	5 23 AEG60659	Polyimmunoglobulin
50	5	100.0	5 23 AAE22204	Peptide linker #2
51	5	100.0	5 23 AAB75959	Linker peptide. S
52	5	100.0	5 23 AAE21141	Linker peptide, to
53	5	100.0	5 23 AAO17155	Human endoglin (CD
54	5	100.0	5 23 AAE19822	Linker peptide #1.
55	5	100.0	5 23 AAE19513	Linker peptide use
56	5	100.0	5 23 AAO14387	Linker peptide for
57	5	100.0	5 23 AAU75495	Human Gly4Ser pept
58	5	100.0	5 23 AAU76043	Synthetic linker m
59	5	100.0	5 23 AAU10629	Antibody variable
60	5	100.0	5 23 AAU74892	Complement pathway
61	5	100.0	5 23 AAU75148	Single unit of lin
62	5	100.0	5 24 AAE34831	Peptide linker #1.
63	5	100.0	5 24 AAE35002	Linker peptide use
64	5	100.0	5 24 AAB72514	Peptide linker. S
65	5	100.0	5 24 AAB75993	Spacer peptide #1.
66	5	100.0	5 24 AEG75221	Zinc finger protei
67	5	100.0	5 24 AAB71088	Amino acid sequenc
68	5	100.0	5 24 AAO26593	Fusion protein rel
69	5	100.0	5 24 AAB98913	Variable region an
70	5	100.0	5 24 AAB56232	Single chain antib
71	5	100.0	5 24 AAE33354	Linker peptide use
72	5	100.0	5 24 AABU5752	Polylinker peptide
73	5	100.0	5 24 AEG72271	Linker peptide #2
74	5	100.0	5 24 AAE30437	Peptide linker #2
75	5	100.0	5 24 AAE30955	Glycine rich pepti
76	5	100.0	5 24 AAO16029	Cyclic single-stra
77	5	100.0	6 15 AAR62168	U1 snRNP 70K prote
78	5	100.0	6 18 AAW17095	Gly(5)-Ser linker
79	5	100.0	6 18 AAW21967	Linker #1 for immu
80	5	100.0	6 18 AAW21967	Linker #1 for immu
81	5	100.0	6 23 ABB78084	ITALY, LOR-2, STRI
82	5	100.0	6 23 ABB84094	Human scdb CEAG1

83 5 100.0 6 23 AAO14388 Linker peptide for
 84 5 100.0 6 23 AAU76044 Synthetic linker m
 85 5 100.0 6 23 AAU82046 T-cell specific bi
 86 5 100.0 6 24 AAE34832 Peptide linker #2.
 87 5 100.0 6 24 ABP71089 Amino acid sequenc
 88 5 100.0 7 17 AAR99245 (Gly6)Ser linker.
 89 5 100.0 7 20 AAY23703 Peptide identified
 90 5 100.0 7 20 AAY02129 Peptide linker use
 91 5 100.0 7 22 ABB56487 Human single chain
 92 5 100.0 7 22 AAB60900 Linker peptide. S
 93 5 100.0 7 23 ABBG32805 Human IGE receptor
 94 5 100.0 7 23 ABBG69318 Synthetic seven am
 95 5 100.0 7 24 AAG74134 Peptide linker ant
 96 5 100.0 7 24 AAE33337 Linker peptide use
 97 5 100.0 7 24 AAE32389 Linker peptide #1
 98 5 100.0 7 24 AAE30034 Peptide used to il
 99 5 100.0 8 17 AAR86795 GM-CSF/SP0 linker
 100 5 100.0 8 20 AAY43498 Linker for dual av
 101 5 100.0 8 21 AAY83212 Peptide linker use
 102 5 100.0 8 22 ABB92194 Peptide hinge used
 103 5 100.0 8 23 ABB97838 Light chain-heavy
 104 5 100.0 8 23 ABB31506 Flexible peptide l
 105 5 100.0 8 24 ABB57935 Zinc finger protel
 106 5 100.0 9 14 AAR31941 In vivo tumour bin
 107 5 100.0 9 16 AAR77978 Conserved Tbp2 pep
 108 5 100.0 9 18 AAW08976 Conserved epitope
 109 5 100.0 9 18 AAW43015 Conserved epitope
 110 5 100.0 9 19 AAW54140 H. influenzae Tbp2
 111 5 100.0 9 19 AAW54141 H. influenzae Tbp2
 112 5 100.0 9 21 AAY51796 H. influenzae tran
 113 5 100.0 9 21 AAY51797 H. influenzae tran
 114 5 100.0 9 21 AAY80384 H. influenzae tran
 115 5 100.0 9 21 AAY80385 H. influenzae tran
 116 5 100.0 9 22 AAG77869 DBA 2 adapter pept
 117 5 100.0 9 22 AAG64069 Hxapha-L9-beta li
 118 5 100.0 9 22 AAB69047 Kylanase related a
 119 5 100.0 9 23 ABBG7839 Light chain-heavy
 120 5 100.0 9 23 ABBG79422 TNFRI helical regi
 121 5 100.0 9 23 ABBG79423 TNFRI helical regi
 122 5 100.0 9 23 AAU76734 Nucleic acid bindi
 123 5 100.0 9 23 AAE13620 Peptide linker #2
 124 5 100.0 9 24 ABBG72089 pUC19-42K spacer/1
 125 5 100.0 10 21 AAB10789 B. subtilis lumazi
 126 5 100.0 10 21 AAY97238 Peptide linker seq
 127 5 100.0 10 22 AAG77866 HLA adapter peptid
 128 5 100.0 10 22 AAE06270 Glycine linker #3
 129 5 100.0 10 22 AAG97432 Human complementar
 130 5 100.0 10 22 AAG97511 Human complementar
 131 5 100.0 10 22 AAG97512 Human complementar
 132 5 100.0 10 22 AAG97513 Human complementar
 133 5 100.0 10 22 AAG97514 Human complementar
 134 5 100.0 10 22 AAG97594 Human complementar
 135 5 100.0 10 22 AAG97626 Human complementar
 136 5 100.0 10 22 AAG97628 Human complementar
 137 5 100.0 10 22 AAG97632 Human complementar
 138 5 100.0 10 22 AAG97634 Human complementar
 139 5 100.0 10 22 AAG97636 Human complementar
 140 5 100.0 10 22 AAG97766 Human complementar
 141 5 100.0 10 22 AAG97768 Human complementar
 142 5 100.0 10 22 AAG97769 Human complementar
 143 5 100.0 10 22 AAG97770 Human complementar
 144 5 100.0 10 22 AAG97788 Human complementar
 145 5 100.0 10 22 AAG97963 Human complementar
 146 5 100.0 10 22 AAG97964 Human complementar
 147 5 100.0 10 22 AAG97994 Human complementar
 148 5 100.0 10 22 AAG83528 Arabidopsis thalia
 149 5 100.0 10 22 AAG83528 Arabidopsis thalia
 150 5 100.0 10 22 AAG83786 Arabidopsis thalia

ALIGNMENTS

RESULT 1
 AAR34034
 ID AAR34034 standard; Protein: 5 AA.
 XX
 AC AAR34034;
 XX
 DT 25-MAR-2003 (updated)
 DT 13-AUG-1993 (first entry)
 XX
 DE Linking sequence whose encoding DNA can be ligated between an
 DE apo A-I- and a B-100-encoding DNA sequence.
 XX
 KW Lipoprotein; apoprotein; B-100; A-I; LDL; HDL; assay.
 XX
 OS Synthetic.
 XX
 PN WO9307165-A1.
 XX
 PD 15-APR-1993.
 XX
 PF 09-OCT-1992; 92WO-US08634.
 XX
 PR 09-OCT-1991; 91US-0774633.
 PR 18-JUN-1992; 92US-0901706.
 PR 08-OCT-1992; 92US-C959946.
 XX
 PA (SCRI) SCRIPPS RES INST.
 XX
 PI Curtiss LK, Koduri KR, Smith RS, Witztum JL, Young SG;
 WPI; 1993-134378/16.
 XX
 DR Polypeptide mimic of native apo B-100 and native apo A-I - useful
 in assays for LDL and HDL in plasma samples
 PT
 XX Disclosure; Page 14 and page 35; 137pp; English.
 PS
 XX The inventors claim a portion of the polypeptide contg. apo B-100
 CC that immunoreacts with antibodies secreted by the hybridoma MB47
 CC having ATCC Accession No. 8746. Polypeptides specifically claimed
 CC include residues 217-297, 216-310, 216-331, 216-352, 216-377, 1-377,
 CC 205-297, 173-297, 140-297. DNA sequences encoding the polypeptides
 CC are also claimed. Also claimed are a fusion polypeptide that
 CC contains: (a) a first amino acid residue sequence up to 250 residues
 CC in length that includes residues 120-135 of apo A-I, (b) a second
 CC amino acid residue sequence up to 375 residues in length that
 CC includes residues 217-297 of apo B-100 and DNA encoding it.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 CC (Updated on 25-MAR-2003 to correct PR field.)
 XX
 SQ Sequence 5 AA;
 Query Match 100.0%; Score 5; DB 14; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GGCGS 5
 Db 1 GGCGS 5
 |||||
 1 GGCGS 5
 RESULT 2
 AAR72707
 ID AAR72707 standard; Peptide; 5 AA.
 XX
 AC AAR72707;
 XX
 DT 31-OCT-1995 (first entry)
 XX
 DE Linker for apo A-I and apo B-100 fusion polypeptide.
 XX
 KW Apo A-I; LDL cholesterol; low density lipoprotein;
 fusion polypeptide; linker.

XX Synthetic.
 XX US5408038-A.
 XX PD 19-APR-1995.
 XX XX 09-OCT-1991; 91US-0774633.
 XX XX 09-OCT-1991; 91US-0774633.
 XX PR 18-JUN-1992; 92US-0901706.
 XX PR 08-OCT-1992; 92US-0959946.
 XX XX (SCRI) SCRIPPS RES INST.
 XX PI Curtiss LK, Koduri KR, Smith RS, Witztum JL, Young SG;
 XX WIPI; 1995-161146/21.
 XX XX New apo.lipoprotein B-100 peptide(s) and fusion peptide(s) - used
 PT in assay systems for detecting LDL and HDL cholesterol levels in
 PT body fluids.
 XX XX Disclosure; Column 18; 41pp; English.
 XX XX A dispersible apo A-I/B-100 fusion polypeptide is claimed which
 CC contains a first AA sequence of apo A-I (see AAR72605) and that
 CC at least AA sequence positions 120-135 (see AAR72606). The two
 CC sequences are operatively linked. An exemplary linking sequence is
 CC AAR72707 whose encoding DNA can be ligated between an apo A-I and a
 CC B-100 encoding DNA sequence.
 XX XX Sequence 5 AA;
 SQ Query Match 100.0%; Score 5; DB 16; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGGS 5
 DB 1 GGGGS 5
 RESULT 3
 ID AAR95062 standard; Peptide; 5 AA.
 AC AAR95062;
 XX 18-AUG-1996 (first entry)
 XX scFv spacer peptide.
 XX Nucleic acid transfer system; gene transfer; gene therapy;
 KW cell targeting; multidomain protein; vector; cancer; scFv;
 KW single chain antibody.
 XX Synthetic.
 XX WO9613599-A1.
 XX PD 09-MAY-1996.
 XX 31-OCT-1995; 95WO-EP04270.
 XX 01-NOV-1994; 94EP-0810627.
 XX (WELS/) WELS W.
 XX Fominaya J, Wels W;
 XX WIPI; 1996-239505/24.
 XX

PT Nucleic acid transfer system for gene therapy, e.g. against cancer
 PT - includes toxin translocation domain to target nucleic acid to
 PT specific cell
 XX Disclosure; Page 8; 106pp; English.
 XX A flexible spacer peptide (AAR95062) is used to link the light chain
 CC variable domain to the heavy chain variable domain of a single
 CC chain recombinant antibody (scFv). The scFv may be derived from
 CC a monoclonal antibody, e.g. MAB FRP5, and forms the ligand domain
 CC of a multidomain protein (see also AAR95053 and AAR95056-58) that is
 CC used with an effector nucleic acid in a novel nucleic acid transfer
 CC system suitable for gene therapy. The ligand domain has a target cell
 CC recognition function and allows cellular internalization of the
 CC multidomain protein/nucleic acid complex.
 XX Sequence 5 AA;
 SQ Query Match 100.0%; Score 5; DB 17; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGGS 5
 DB 1 GGGGS 5
 RESULT 4
 ID AAR17094 standard; peptide; 5 AA.
 XX AC AAR17094;
 XX 14-SEP-1999 (first entry)
 XX Gly(4)-Ser linker peptide for chimeric protein construct.
 DE Haematopoietic protein; human; granulocyte-colony stimulating factor;
 KW G-CSF; interleukin; c-mpl ligand; linker; gene therapy; aplastic anaemia;
 KW stem cell expansion; leukopaenia; neutropaenia; vector; bone marrow;
 KW thrombocytopaenia; blood cell activation; growth.
 XX Synthetic.
 XX WO9712985-A2.
 XX 10-APR-1997.
 XX 04-OCT-1996; 96WO-US15774.
 XX 05-OCT-1995; 95US-0004834.
 XX (SEAR) SEARLE & CO G D.
 XX Bauer SC, Baum CM, Caparon MH, Feng Y, Giri JG;
 PI Klein BK, Lee SC, McKearn JP, McWherter CA, Staten NR;
 PI Summers NL, Zurfluh L;
 XX WIPI; 1997-226228/20.
 XX Multi-functional haematopoietic receptor agonists - used to
 PT stimulate the production of haematopoietic cells in patients
 XX Disclosure; Page 33; 616pp; English.
 XX The invention relates to a novel haematopoietic protein (HP) comprising
 CC an amino acid (AA) sequence of formula: R1-L1-R2; R2-L1-R1; R1-R2; or
 CC R2-R1; where R1 and R2 are independently selected from: (I) a modified
 CC human granulocyte-colony stimulating factor (hG-CSF) AA sequence;
 CC (II) a modified human interleukin-3 (hIL-3) AA sequence; (III) a modified
 CC human c-mpl ligand; and a colony stimulating factor (CSF); and L1 = a
 CC linker capable of linking R1 to R2. This sequence represents an example
 CC of a linker used to construct the proteins of the invention.

C Vectors comprising the nucleic acid molecules are useful for the recombinant production of HP. The nucleic acid molecules are useful in gene therapy. The HP's are useful for stimulating the production of haematopoietic cells in patients, selective ex vivo expansion of stem cells and for treatment of haematopoietic disorders. Disorders that can be treated include leukopenia, neutropenia, aplastic anaemia and thrombocytopenia. In vitro uses include the ability to stimulate bone marrow and blood cell activation and growth before infusion into the patients.

CC patients.

CC Sequence 5 AA;

Query Match 100.0%; Score 5; DB 18; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5

DB 1 GGGGS 5

RESULT 5

AAW19543
AAW19543 standard; peptide; 5 AA.

AC AAW19543;

DT 19-FEB-1998 (first entry)

XX Chimeric protein pentapeptide linker for the MBP moiety and PE moiety.

XX Pseudomonas exotoxin; myelin basic protein; chimeric protein;
XX autoimmune disease; multiple sclerosis; human.

XX Synthetic.

XX WO9719179-A1.

XX 29-MAY-1997.

XX 17-NOV-1996; 96WO-IL00151.

XX 26-DEC-1995; 95IL-0116559.

XX 17-NOV-1995; 95IL-0116044.

XX (YISS) YISSUM RES & DEV CO.

XX Beraud E, Lorberboun-Galski H, Marianovsky I, Steinberger I;
XX Yarkoni S;

XX WPI; 1997-298116/27.

XX New Pseudomonas exotoxin-myeelin basic protein chimeric proteins -
XX used for the treatment of autoimmune diseases, particularly
XX multiple sclerosis

XX Claim 6; Page 29; 54pp; English.

XX New chimeric proteins have been developed comprising a Pseudomonas aeruginosa exotoxin (PE) moiety linked to a myelin basic protein (MBP) moiety selected from: (a) MBP; (b) amino acids 69-88 of guinea-pig MBP or an antigenic portion; (c) amino acids 84-102 of human MBP or an antigenic portion; (d) amino acids 143-168 of human MBP or an antigenic portion; and (e) an amino acid sequence in which one or more amino acids have been deleted, added, substituted or mutated in the amino acid sequences of (a), (b), (c), or (d), the modified sequences retaining at least 75% homology with the amino acid sequences. The present sequence represents the preferred pentapeptide linker used to link the MBP moiety and PE moiety in a chimeric protein. The chimeric proteins can be used for the treatment of autoimmune diseases such as multiple sclerosis. The chimeric proteins can specifically target and kill MBP specific T cells while having no effect on non-target cells.

XX Sequence 5 AA;

Query Match 100.0%; Score 5; DB 18; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5

DB 1 GGGGS 5

RESULT 6

AAV43496

ID AAV43496 standard; Peptide; 5 AA.

AC AAV43496;

DT 26-JAN-2000 (first entry)

XX Linker for dual avb3 receptor/metastasis-associated receptor ligands.

XX Interferon-alpha-2b; IFN-alpha; avb3 antagonist; avb3 receptor ligand;
XX metastasis-associated receptor ligand; angiogenesis; cell proliferation;
XX anti-angiogenic protein; avb3-integrin; cancer; arthritis;

XX macular degeneration; diabetic retinopathy; hemangioma; psoriasis;

XX osteoporosis; thrombosis; angina; atherosclerosis; antiviral;

XX antibacterial; antifungal.

XX Homo sapiens.

XX WO9951638-A1.

XX 14-OCT-1999.

XX 07-APR-1999; 99WO-US04295.

XX 08-APR-1998; 98US-0081074.

XX (SEAR) SEARLE & CO G D.

XX Tjoeng FS, Fok KF;

XX WPI; 1999-620196/53.

XX New conjugates of integrin antagonist and ligand for
XX metastasis-associated receptor, for treating angiogenesis-related
XX diseases, e.g. cancer

XX Claim 18; Page 86; 108pp; English.

XX The present sequence represents a linker used to join the avb3 antagonist and the metastasis-associated receptor ligand, in the pharmaceutical compounds of the invention. These compounds are dual avb3 receptor/metastasis-associated receptor ligands, and inhibit angiogenesis and thus proliferation of (cancer) cells. One component binds to the avb3 receptor and the other to a metastasis-associated receptor. The avb3 antagonists may also be conjugated to anti-angiogenic proteins, such as IFN-alpha and its derivatives. The compounds are used to treat angiogenesis-related disorders (mediated by the avb3-integrin), specifically cancer (of lung, breast, ovary, prostate, stomach, colon, kidney or bladder, also melanoma, hepatoma, sarcoma and lymphoma), arthritis and macular degeneration, and also diabetic retinopathy, hemangioma, psoriasis, osteoporosis, thrombosis, angina, atherosclerosis etc. The compounds may also be useful as antiviral, antibacterial and antifungal agents.

XX Sequence 5 AA;

Query Match 100.0%; Score 5; DB 20; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5

|||||


```

Db      1 GGGGS 5

RESULT 7
AAV33597
ID AAY33597 standard; Protein; 5 AA.
XX
AC AAY33597;
XX
DT 20-DEC-1999 (first entry)
XX
DE VH-VL domain linker peptide #9.
XX
KW Antigen binding; single chain; variable domain; VH domain; light chain;
KW heavy immunoglobulin chain; VL domain; anticancer; antiviral; tumor;
KW antibacterial; antimalarial; antiinflammatory; treatment; prevention;
KW diagnosis; vaccine; autoimmune disease; inflammation; blood disorder;
KW transplant rejection; arthritis; nervous system disorder; infection.
XX
OS Synthetic.
XX
PN DE19816141-A1.
XX
PD 14-OCT-1999.
XX
PF 09-APR-1998; 98DE-1016141.
XX
PR 09-APR-1998; 98DE-1016141.
XX
PA (HME1) HOECHST MARION ROUSSEL DEUT GMBH.
XX
PI Kontermann R, Sedlacek H, Mueller R;
XX
DR WPI; 1999-581511/50.
XX
PT New polyclonal binding agents containing variable heavy and light
PT constructs connected via peptide linker, used for treatment, prevention
PT or diagnosis of e.g. cancer -
XX
XX
PS Claim 7; Page 17; 20pp; German.
XX
CC This sequence represents a novel single-chain molecule (I) that binds
CC multiple antigens and comprises two variable domains of heavy
CC immunoglobulin chains (VH), having specificities A and B and two
CC variable domains of light chains (VL), also with specificities A and B.
CC The domains are provided as two VH-VL constructs which are attached via
CC a peptide (P). Any VH and VL may be replaced by their functional
CC fragments. The products of the invention have anticancer, antiviral,
CC antibacterial, antimalarial and antiinflammatory activity. (I) are used
CC to treat, prevent or diagnose tumors (e.g. as tumor vaccines), autoimmune
CC diseases and inflammation (e.g. transplant rejection and arthritis),
CC blood disorders (e.g. of the coagulation and/or circulatory systems, such
CC as anemia, leucopenia, thrombocytopenia and hypertension), nervous system
CC disorders and/or infections (by viruses or bacteria, or malaria),
CC including, when (I) include a fusogenic peptide, use for gene transfer.
CC (I) are produced simply and in predominantly homogeneous form, in a wide
CC variety of hosts, either in secreted or membrane-bound forms. This
CC sequence represents a VH-VL domain linker peptide which is used to
CC illustrate the method of the invention.
XX
SQ Sequence 5 AA;

Query Match 100.0%; Score 5; DB 20; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GGGGS 5
Db 1 GGGGS 5

RESULT 8
AAV25357
ID AAY25357 standard; Protein; 5 AA.
XX
AC AAY25357;
XX
DT 06-SEP-1999 (first entry)
XX
DE IFNAR2/IFN-beta complex peptide fragment 1.
XX
KW IFNAR2; IFN-beta; type I interferon; IFNAR/IFN complex; IFN; antiviral;
KW human interferon alpha/beta receptor; anticancer; immunomodulatory;
KW anti-arthritis; antidiabetic; treatment; hepatitis; viral infection;
KW hairy cell leukemia; Kaposi's sarcoma; multiple myeloma; cancer; lupus;
KW diabetes; multiple sclerosis; rheumatoid arthritis; myasthenia gravis;
KW acquired immune deficiency syndrome.
XX
OS Synthetic.
XX
PN WO9932141-A1.
XX
PD 01-JUL-1999.
XX
PF 18-DEC-1998; 98WO-US26926.
XX
PR 19-DEC-1997; 97US-0068295.
XX
PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
XX
PI (MCIN/) MCINNIS P G.
XX
DR Cunningham M, El Tayar N, McKenna S, Sherris D;
XX
DR WPI; 1999-405115/34.
XX
PT Prolonging in vivo activity of type I interferon by complexing
XX Example 8; Page 76; 86pp; English.
XX
CC This invention describes a novel method for prolonging the in vivo effect
CC of type I interferon (IFN) by administering IFN as a complex (A) with a
CC subunit (I) of the human interferon alpha/beta receptor (IFNAR). The
CC product of the invention has antiviral, anticancer, immunomodulatory,
CC anti-arthritis and antidiabetic activity. (A) have the antiviral,
CC anticancer and immunomodulating activities of IFN, e.g. for treating
CC hepatitis and other viral infections, hairy cell leukemia, Kaposi's
CC sarcoma, multiple myeloma and other cancers, multiple sclerosis,
CC rheumatoid arthritis, myasthenia gravis, diabetes, acquired immune
CC deficiency syndrome and lupus. When complexed in (A), the storage life of
CC IFN is increased (i.e. it is stabilized against oligomerization, without
CC the need for storage at acidic pH) and its biological effect is
CC potentiated.
XX
SQ Sequence 5 AA;

Query Match 100.0%; Score 5; DB 20; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GGGGS 5
Db 1 GGGGS 5

RESULT 9
AAV02127
ID AAY02127 standard; Protein; 5 AA.
XX
AC AAY02127;
XX
DT 16-JUL-1999 (first entry)
XX
DE Peptide linker used to make multifunctional proteins.
XX
KW Angiostatin; endostatin; interferon; thrombospondin;

```

interferon-inducible protein; platelet factor 4; anti-angiogenic; anti-tumor; multifunctional protein; angiogenic-mediated disease; cancer; diabetic retinopathy; macular degeneration; arthritis; tumor cell production; peptide linker.

Homo sapiens.

W09916889-A1.

08-APR-1999.

30-SEP-1998; 98WO-US20464.

01-OCT-1997; 97US-0060609.

(SEAR) SEARLE & CO G D.

Bolanowski MA, Caparon MH, Casperson GF, Gregory SA; Klein BK, McKearn JP;

WPI; 1999-255098/21.

New multifunctional proteins useful for treating angiogenic-mediated diseases

Disclosure; Page 111; 121pp; English.

The specification describes multifunctional proteins which comprise combinations of angiotensin, endostatin, interferon, thrombospondin, interferon-inducible protein and platelet factor 4, and have anti-angiogenic and/or anti-tumor activity. The multifunctional protein may exhibit useful properties such as having similar or greater biological activity when compared to a single factor or by having improved half-life or decreased adverse side effects, or a combination of these properties. The proteins can be used for treating an angiogenic-mediated disease, e.g. cancer, diabetic retinopathy, macular degeneration, or arthritis. They can also be used for inhibiting the production of tumor cells (characteristic of lung, breast, ovarian, prostate, pancreatic, gastric, colon, renal, bladder cancers; melanoma, hepatoma, sarcoma and lymphoma) in a patient and for inhibiting tumor growth. AAR0125-32 represent peptide linkers used to make the multifunctional proteins of the invention.

Sequence 5 AA;

Query Match 100.0%; Score 5; DB 20; Length 5;

Best Local Similarity 100.0%; Pred. No. 9.3e+05; Mismatches 0; Indels 0; Gaps 0;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5

|||||

1 GGGGS 5

Db

RESULT 10

AAB00156

AAB00156 standard; Peptide; 5 AA.

AAB00156;

08-FEB-2001 (first entry)

Linker used in sCD4-SCFv(17b) fusion protein.

Fusion protein; HIV; human immunodeficiency virus; antibody; Fv;

AIDS; acquired immune deficiency syndrome; neutralisation;

infection; gene therapy; CD4; gp120; glycoprotein; resistance;

vaccination; binding domain; single chain antibody; chimera;

chimeric protein.

Synthetic.

OS XX

W0200055207-A1.

XX 21-SEP-2000.

PD 16-MAR-2000; 2000WO-US06946.

XX 16-MAR-1999; 99US-0124681.

PR (USSH) US NAT INST OF HEALTH.

XX Berger EA, Del Castillo CM;

PA WPI; 2000-638183/61.

PI Novel neutralizing bispecific fusion proteins effective in viral such

XX as HIV neutralization, comprises two different binding domains, linked

PT inducing-binding domain and induced-binding domain functionally linked

PT by linker

PS Claim 30; Page 45; 55pp; English.

XX sCD4-SCFv(17b) is a neutralising bispecific fusion protein capable of

CC binding to two sites of its target protein. The protein comprises a

CC first binding domain capable of binding to an inducing site on the

CC target protein, a second binding domain capable of forming

CC neutralising complex with an induced epitope of the target protein and

CC a linker connecting the binding domains. sCD4-SCFv(17b) comprises a

CC soluble CD4 fragment (containing domains D1 and D2) fused to a single

CC chain Fv portion of antibody 17b via a linker. sCD4-SCFv(17b), its

CC variant, analogue or mimetic is used for inactivating gp120 protein of

CC HIV, and for neutralising HIV. It is also used for blocking and

CC preventing the binding of the viral or recombinant gp120 protein to

CC soluble CD4 or lymphocyte CD4 and for inhibiting HIV replication. The

CC chimeric proteins is therefore useful for treating HIV infection and

CC also AIDS. It is particularly useful in the prevention of

CC infection during or immediately after HIV exposure (e.g.,

CC mother/infant transmission, post-exposure prophylaxis, and as a

CC topical inhibitor) and for providing long term resistance to HIV

CC infections and AIDS. Gene therapy is used to secrete the bispecific

CC protein at mucosal surfaces, such as the vaginal, rectal or oral

CC mucosa. The fusion proteins is highly potent, broadly cross-reactive

CC with neutralising antibody with high in vivo activity and no

CC FC-mediated undesirable targetting properties. When the fusion

CC protein is substantially derived from human proteins, it has minimal

CC immunogenicity and toxicity in humans which is of great value in

CC prevention of infection during or immediately after HIV exposure.

XX Query Match 100.0%; Score 5; DB 21; Length 5;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5

|||||

1 GGGGS 5

Db

RESULT 11

AAB14535

ID AAB14535 standard; peptide; 5 AA.

XX AAB14535;

AC 24-NOV-2000 (first entry)

XX Peptide linker for joining HIV-1 gp41 N- and C-terminal helices.

DE HIV-1; gp41; N-helical domain; heptad repeat region; C-helical domain;

XX gp41 transmembrane-proximal amphipathic alpha-helical segment;

XX core 6-helix bundle; viral entry inhibition; immunogenic;

XX antibody; humoral response; broad spectrum vaccine; anti-HIV;

XX envelope glycoprotein; prophylaxis; therapy; peptide linker.

XX

OS Synthetic.
 XX WO200040616-A1.
 XX 13-JUL-2000.
 XX 10-JAN-2000; 2000WO-US00456.
 XX 08-JAN-1999; 99US-C115404.
 XX 07-JAN-2000; 2000US-0480336.
 XX (WILD/) WILD C T.
 PA (WEIS/) WEISS C D.
 XX Wild CT, Weiss CD;
 XX WPI; 2000-465959/40.
 XX Raising neutralizing antibody response to human immunodeficiency virus,
 PT comprises administering a polypeptide capable of forming a stable
 PT coiled-coil solution structure.
 XX Disclosure; Page 15; 97pp; English.
 XX The invention relates to raising a neutralising antibody response to a
 CC broad spectrum of HIV (human immunodeficiency virus) strains and
 CC isolates, comprising the administration of a peptide which corresponds
 CC to or mimics highly conserved portions of the gp41 envelope glycoprotein
 CC which are important in mediating the process of viral entry into host
 CC cells. Such peptides can correspond to or mimic the coiled coil
 CC solution structure of the N-helical domain (the heptad repeat
 CC region), or can correspond or mimic the C-helical domain (the
 CC transmembrane-proximal amphipathic alpha-helical segment), or the
 CC gp41 core 6-helix bundle, which is formed by the interaction of
 CC the N- and C-helical domains of three gp41 proteins. The peptides
 CC can be administered either singly or as a combination (particularly
 CC a combination of N-helical and C-helical peptides), and can be
 CC multimerised. For example, N- and C-helical domain peptides can be
 CC alternately linked together to form a peptide which mimics the core
 CC 6-helix bundle. Administration of the peptide(s) generates a humoral
 CC response, with the production of antibodies against gp41 structures
 CC involved in viral entry. As these portions of gp41 are well conserved,
 CC such antibodies may be effective against a broad range of HIV strains
 CC and isolates. The peptide compositions may be administered as a
 CC prophylactic or therapeutic vaccine to generate antibodies which reduce
 CC or inhibit the ability of HIV to infect uninfected cells. A composition
 CC comprising polyclonal or monoclonal antibodies can be administered to
 CC reduce HIV infection of uninfected cells. Antibodies raised against
 CC entry-relevant gp41 structures may also be used therapeutically and as
 CC tools to further elucidate the mechanism of HIV cell entry. The
 CC present sequence represents a peptide linker which may be used to
 CC join peptides of the invention together to form multimers.
 XX Sequence 5 AA;
 SQ Query Match 100.0%; Score 5; DB 21; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGGS 5
 Db |||||
 1 GGGGS 5
 RESULT 13
 ID AAY83210 standard; Peptide; 5 AA.
 XX AAY83210;
 AC AAY83210;
 XX 24-JUL-2000 (first entry)
 DT Peptide linker used in construction of a_vb_3 integrin/IFN alpha.
 XX Biconjugate; a_vb_3 integrin; interferon alpha; angiogenesis;
 XX cancer; tumour; osteoporosis; Paget's disease; Kaposi's sarcoma;
 XX periodontal disease; metastasis; neoplasia; retinopathy; arthritis;
 XX psoriasis; leukaemia; malignant melanoma; atherosclerosis;
 XX smooth muscle cell migration; inhibition; treatment; antagonist;
 XX angina; thrombosis; restenosis; antiviral; antifungal;
 XX antibacterial.
 OS Synthetic.
 XX WO200009143-A1.
 XX 24-FEB-2000.
 XX 07-APR-1999; 99WO-US04296.
 XX 13-AUG-1998; 98US-0096442.
 DE Expression vector CANTAB5E inserted peptide.

XX Modified Rhase; eosinophil derived neurotoxin protein; EDN; cancer;
 KW Kaposi's sarcoma; neoplastic endothelial cell;
 KW non-neoplastic endothelial cell; expression vector.
 XX Synthetic.
 OS WO200026233-A1.
 XX 11-MAY-2000.
 XX 01-NOV-1999; 99WO-US25737.
 XX 02-NOV-1998; 98US-0106732.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Rybak SM, Newton DL;
 PI WPI; 2000-365565/31.
 XX N-terminally modified Rhase A targeted to and are cytotoxic to
 PT cancerous endothelial cells used to treat especially Kaposi's sarcoma
 PT
 XX Example 9; Page 34; 51pp; English.
 XX The present sequence is a peptide which was inserted into expression
 CC vector pCANTAB5E to enable more flexible folding of the human eosinophil
 CC derived neurotoxin protein (EDN), which was expressed by the vector. This
 CC protein can be directed to cancerous cells using additional N-terminal
 CC peptides, where it exerts a cytotoxic effect. The protein can, therefore,
 CC be used to treat cancer, particularly Kaposi's sarcoma, and to
 CC selectively kill neoplastic and non-neoplastic endothelial cells.
 XX Sequence 5 AA;
 SQ Query Match 100.0%; Score 5; DB 21; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGGS 5
 Db |||||
 1 GGGGS 5
 RESULT 13
 ID AAY83210 standard; Peptide; 5 AA.
 XX AAY83210;
 AC AAY83210;
 XX 24-JUL-2000 (first entry)
 DT Peptide linker used in construction of a_vb_3 integrin/IFN alpha.
 XX Biconjugate; a_vb_3 integrin; interferon alpha; angiogenesis;
 XX cancer; tumour; osteoporosis; Paget's disease; Kaposi's sarcoma;
 XX periodontal disease; metastasis; neoplasia; retinopathy; arthritis;
 XX psoriasis; leukaemia; malignant melanoma; atherosclerosis;
 XX smooth muscle cell migration; inhibition; treatment; antagonist;
 XX angina; thrombosis; restenosis; antiviral; antifungal;
 XX antibacterial.
 OS Synthetic.
 XX WO200009143-A1.
 XX 24-FEB-2000.
 XX 07-APR-1999; 99WO-US04296.
 XX 13-AUG-1998; 98US-0096442.

XX (SEAR) SEARLE & CO G D.
 XX
 XX Fok KF, Tjoeng FS;
 XX WPI; 2000-205894/18.
 XX
 XX New bioconjugates comprising an avb3 antagonist and a
 XX metastatic-associated receptor ligand, useful for treating cancer and
 XX other angiogenic diseases, or as antiviral, antifungal or antibacterial
 XX agents
 XX
 XX Claim 19; Page 88; 123pp; English.
 XX
 XX Bioconjugates comprising one or more a vb3 antagonist moieties
 XX coupled to a peptide or polypeptide having anti-angiogenic properties
 XX can be used for treating a human patient with an
 XX angiogenesis-mediated disease, e.g. cancer, arthritis, or macular
 XX degeneration. The a vb3 integrin is normally associated with
 XX endothelial cells but can promote the formation of blood vessels
 XX (angiogenesis) in tumours. The a vb3 integrin is also known to
 XX play a role in tumour metastasis, neoplasia, osteoporosis,
 XX Paget's disease, retinopathy, arthritis, periodontal disease,
 XX psoriasis and smooth muscle cell migration. Interferon alpha is a
 XX family of proteins which possess complex antiviral, antineoplastic
 XX and immunomodulating activities. Interferon alpha is effective
 XX against a variety of cancers including hairy cell leukaemia,
 XX chronic myelogenous leukaemia, malignant melanoma and Kaposi's
 XX sarcoma. Multi-functional bioconjugates comprising both a vb3
 XX antagonist and interferon alpha 2b can exhibit greater biological
 XX activity when compared to a single factor or having improved
 XX half-life or decreased adverse side effects, or a combination of
 XX these properties. They can be used for inhibiting elevated levels
 XX of tumor antigens, inhibiting the proliferation of tumor cells and
 XX inhibiting tumor growth. The bioconjugates can also be used for
 XX treating e.g. osteoporosis, humoral hypercalcemia of malignancy,
 XX Paget's disease, retinopathy including diabetic retinopathy,
 XX arthritis, including rheumatoid arthritis, periodontal disease,
 XX psoriasis, thrombosis, angina, atherosclerosis, smooth muscle cell
 XX migration and restenosis in a mammal. They are also useful as a
 XX antiviral, antifungal and antibacterial agents. This sequence is a
 XX peptide linker used in the construction of the multi-functional
 XX bioconjugates.
 XX
 XX Sequence 5 AA;
 XX
 XX Query Match 100.0%; Score 5; DB 21; Length 5;
 XX Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 1 GGGGS 5
 XX |||||
 XX Db 1 GGGGS 5
 XX
 XX RESULT 14
 XX ID AAY43750
 XX AAY43750 standard; Peptide; 5 AA.
 XX
 XX AC AAY43750;
 XX
 XX DT 11-FEB-2000 (first entry)
 XX
 XX Linker used to construct a bispecific single-chain antibody.
 XX
 XX bscCD19xCD3 antibody; bispecific single-chain fragment; CD19 antigen;
 XX CD3 antigen; CD19-positive target cell; T-cell stimulation;
 XX cytotoxic T-lymphocyte; B-cell malignancy; myasthenia gravis;
 XX B-cell mediated autoimmune disease; Morbus Basedow;
 XX Hashimoto thyroiditis; Goodpasture syndrome; B-cell depletion;
 XX non-Hodgkin lymphoma; gene therapy; cancer; viral disease.
 XX
 XX Synthetic.
 XX

XX WO9954440-A1.
 XX
 XX 28-OCT-1999.
 XX
 XX 21-APR-1999; 99WO-EP02693.
 XX
 XX 21-APR-1998; 98EP-0107269.
 XX
 XX (DOER/) DOERKEN B.
 XX (RIET/) RIETHMUELLER G.
 XX
 XX Kufer P, Lutterbuese R, Bargou R, Loeffler A;
 XX WPI; 2000-013241/01.
 XX
 XX Novel multifunctional polypeptide for treating B-cell malignancies
 XX especially non-Hodgkin lymphoma -
 XX
 XX Claim 10; Page 49; 91pp; English.
 XX
 XX The present sequence represents a linker used in the construction
 XX of bispecific single-chain polypeptides of the invention. These
 XX polypeptides comprise domains providing binding-site of immunoglobulin
 XX chains or antibodies specifically recognizing CD19 and CD3 antigen.
 XX The polypeptide destroys CD19-positive target cells without any need
 XX of T-cell pre and/or co-stimulation, by recruiting cytotoxic
 XX T-lymphocytes and so specific lysis by T-cells rather than a direct
 XX effect by an antibody is achieved. The bispecific single-chain
 XX polypeptides, or nucleotides encoding them, are used for the treatment
 XX of B-cell malignancies, B-cell mediated autoimmune diseases like
 XX myasthenia gravis, Morbus Basedow, Hashimoto thyroiditis or Goodpasture
 XX syndrome or for the depletion of B-cells and more particularly
 XX non-Hodgkin lymphoma in mammals preferably human. They can also delay
 XX the pathological conditions caused by these diseases, and can be used
 XX for detecting these diseases. The polynucleotide is used for gene
 XX therapy. The polypeptides are also used for identifying compounds
 XX modulating B-cell/T-cell mediated immune response with car in turn be
 XX used for treating cancer, its related diseases and also for inhibiting
 XX viral diseases by preventing viral infection.
 XX
 XX Sequence 5 AA;
 XX
 XX Query Match 100.0%; Score 5; DB 21; Length 5;
 XX Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 1 GGGGS 5
 XX |||||
 XX Db 1 GGGGS 5
 XX
 XX RESULT 15
 XX ID AAY54917
 XX AAY54917 standard; peptide; 5 AA.
 XX
 XX AC AAY54917;
 XX
 XX DT 14-FEB-2000 (first entry)
 XX
 XX Linker from IL-12 fusion protein.
 XX
 XX Interleukin-12; IL-12; fusion protein; IL-12 p35 subunit; B7 protein;
 XX IL-12 p40 subunit; gene therapy; tumour; leukaemia; linker.
 XX
 XX Synthetic.
 XX
 XX US5994104-A.
 XX
 XX 30-NOV-1999.
 XX
 XX 08-NOV-1996; 96US-0751767.
 XX

PR 08-NOV-1996; 96US-0751767.
 XX (UNLO) ROYAL FREE HOSPITAL SCHOOL MED.
 PA
 XX Anderson RJ, Prentice HG, MacDonald ID;
 PI
 XX WPI; 2000-038261/03.
 DR
 XX Nucleic acid constructs encoding interleukin-12 fusion proteins useful
 XX for treating leukemia and other cancers -
 PT
 XX Claim 3; Column 93; 73pp; English.
 PS
 XX This sequence represents a linker that can be used in an interleukin-12
 CC fusion protein. The invention relates to an isolated nucleic acid
 CC construct (I) comprising a region encoding an interleukin-12 (IL-12)
 CC fusion protein (comprising an IL-12 p35 subunit, an IL-12 p40 subunit and
 CC a linker peptide (joining the subunits)) and a region encoding a B7
 CC protein. (I) may be used to produce IL-12 fusion proteins according to
 CC standard recombinant DNA methodologies. The fusion proteins may be
 CC produced either in vitro in a fermentation culture or in vivo as part of
 CC a gene therapy protocol (in this case (i) is used to transform a patient's
 CC cells, which then secrete the functional polypeptide to supplement the
 CC patient's own production of IL-12 or to rectify mutations which lead to
 CC the expression of inactive polypeptides). The fusion proteins produced in
 CC this way may be used to treat any disease which responds to IL-12 such as
 CC tumours (both solid and dispersed of the kidney, breast, colon, ovarian
 CC and cervical tumours and melanomas) and in particular, tumours of the
 CC blood such as leukaemia. Alternatively, the polypeptides may be used as
 CC antigens in the production of antibodies to IL-12 and to assay for
 CC agonists and antagonists of its activity. The antibodies and antagonists
 CC may be used to inhibit the activity of IL-12. (I) may also be used
 CC diagnostically as a probe which hybridizes to sequences encoding IL-12
 CC and the antibodies may be used to detect the presence of IL-12
 CC polypeptides in samples. They may be used diagnostically to quantitate
 CC the expression of the polypeptide by patients and hence which subjects
 CC may be in need of restorative therapy.
 XX
 SQ Sequence 5 AA;
 Query Match 100.0%; Score 5; DB 21; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GGGGS 5
 Db |||||
 1 GGGGS 5
 RESULT 16
 ABB56480
 ID ABB56480 standard; Peptide; 5 AA.
 XX
 AC ABB56480;
 XX
 DT 25-FEB-2002 (first entry)
 XX
 DE Human single chain MHC class II molecule linker #4.
 XX
 KW Human; MHC; major histocompatibility complex; MHC class II; multimer;
 KW single chain; immunosuppressive; antidiabetic; antiinflammatory;
 KW antianemic; antirheumatoid; antiarthritic; neuroprotective; vaccine;
 KW autoimmune disease; insulin dependent diabetes; multiple sclerosis;
 KW myasthenia gravis; pernicious anaemia; autoimmune encephalomyelitis;
 KW rheumatoid arthritis; systemic lupus erythematosus; linker.
 XX
 OS Homo sapiens.
 XX
 PN WO200170245-A1.
 XX
 PD 27-SEP-2001.
 XX
 PF 22-MAR-2001; 2001WO-US09616.

XX 22-MAR-2000; 2000US-191274P.
 PR 15-MAY-2000; 2000US-204249P.
 PR 23-JAN-2001; 2001US-264003P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Carter D, Zhu S, Arimilli S, Wang A;
 XX WPI; 2001-616371/71.
 DR
 XX Multimeric complex for treating autoimmune diseases, comprises first
 PT and second single chain MHC class II molecules, each comprising alpha
 PT and beta domain linked through amino acid linker and multimerization
 PT domain -
 XX
 PS Claim 24; Page 131; 147pp; English.
 XX
 CC The invention relates to a multimeric complex comprising a first
 CC recombinant single chain major histocompatibility complex (MHC) class
 CC II molecule and a second recombinant single chain MHC class II molecule,
 CC each comprising an alpha domain and a beta domain linked through an
 CC amino acid linker and a multimerization domain. The first and the
 CC second molecule are linked through the multimerization domain to form
 CC a multimeric complex. The complex is useful for treating autoimmune
 CC diseases. It is useful for treating insulin dependent diabetes,
 CC multiple sclerosis, myasthenia gravis, pernicious anaemia, autoimmune
 CC encephalomyelitis (EAE), rheumatoid arthritis and systemic lupus
 CC erythematosus. The present sequence is a linker used to connect the
 CC alpha and beta domains of the single chain MHC class II molecules of
 CC the invention.
 XX
 SQ Sequence 5 AA;
 Query Match 100.0%; Score 5; DB 22; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GGGGS 5
 Db |||||
 1 GGGGS 5
 RESULT 17
 AAE11873
 ID AAE11873 standard; peptide; 5 AA.
 XX
 AC AAE11873;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Spacer peptide used in the antibody display and screening methods.
 XX
 KW Shuffling; recombinant polynucleotide; selecting; screening; IL-1beta;
 KW functional property; molecular backcross; spacer peptide.
 XX
 OS Unidentified.
 XX
 PN US6277638-B1.
 XX
 PD 21-AUG-2001.
 XX
 PF 15-JAN-1999; 99US-0232863.
 XX
 PR 19-JUN-1998; 98US-0100856.
 PR 17-FEB-1995; 95WO-US02126.
 PR 04-MAR-1996; 96US-0537874.
 PR 17-FEB-1994; 94US-0198431.
 XX
 PA (MAXY-) MAXYGEN INC.
 XX
 XX Stemmer WFC;
 XX

DR WPI; 2001-610489/70.

XX Evolving polynucleotides with desired functional property, involves

PT shuffling polynucleotides comprising species variants to form a

PT population of recombinant polynucleotides, and selecting/screening the

PT population -

XX Disclosure; Column 12; 62pp; English.

PS The present invention relates to a method for evolving a polynucleotide

CC toward a desired functional property. The method comprising providing

CC polynucleotides comprising two or more species variants, shuffling the

CC polynucleotides to form a population of recombinant polynucleotides,

CC selecting or screening for recombinant polynucleotides that have evolved

CC toward the desired functional property, and repeating the steps to obtain

CC the polynucleotide. The methods are useful for evolving a polynucleotide

CC toward a desired functional property and for obtaining a recombinant

CC polynucleotide having a desired functional property. The methods are also

CC useful for generating and selecting nucleic acid fragments encoding

CC mutant proteins. The methods are also useful in molecular backcross to

CC eliminate unnecessary or silent mutations, in the study of repetitive DNA

CC sequences and to mutate ribozymes and aptamers. The present sequence is

CC spacer peptide used in the antibody display and screening methods.

XX SQ Sequence 5 AA;

2y Query Match 100.0%; Score 5; DB 22; Length 5;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GGGGS 5

1 GGGGS 5

RESULT 18

AAE08730

ID AAE08730 standard; peptide; 5 AA.

AC AAE08730;

XX 15-NOV-2001 (first entry)

PT Peptide linker #9 for fusion of zinc finger proteins (ZFP).

XX Molecular target; zinc finger protein; ZFP; cellular process;

GW signal transduction; drug-screening.

XX Synthetic.

CS WO200159450-A2.

PN 16-AUG-2001.

PD 08-FEB-2001; 2001WO-US04301.

XX 08-FEB-2000; 2000US-0181117.

XX (SANG-) SANGAMO BIOSCIENCES INC.

PA Case C;

PI WPI; 2001-522491/57.

XX Screening compound for interaction with molecular target by contacting

PT compound with cells, comprising exogenous zinc finger protein that

PT modulates expression of target, and determining values of properties of

PT cells -

XX Disclosure; Page 28; 99pp; English.

PS The invention relates to a method of screening a compound for interaction

CC with a molecular target. The method involves contacting first and

CC second cells with the compound and determining the values of properties

CC of the compound. The second cell comprises an exogenous zinc finger

CC protein (ZFP) that modulates the expression of the molecular target, or

CC isolating membranes from cell comprising ZFP. The methods allow for high

CC throughput screening of candidate compound and reduces the incidence of

CC false positives. The methods are useful for screening a compound for

CC interaction with a molecular target or for screening a compound for its

CC effect on a cellular process. The method is useful for testing a compound

CC for its capacity to transduce a signal to the molecular target or its

CC capacity to block transduction of a signal through the molecular target,

CC and for performing biochemical drug-screening assays. The present

CC sequence is a peptide linker used for linking ZFPs.

XX SQ Sequence 5 AA;

Query Match 100.0%; Score 5; DB 22; Length 5;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5

Db 1 GGGGS 5

RESULT 19

AAU05172

ID AAU05172 standard; peptide; 5 AA.

AC AAU05172;

XX 24-OCT-2001 (first entry)

DT Pain-relieving target construct, linker moiety #1.

DE Pain relief; tumour; therapeutic; breast cancer; lung cancer;

XX Hodgkin's lymphoma; neuroblastoma; ischaemic.

KW Synthetic.

OS WO200147512-A2.

XX 05-JUL-2001.

XX 06-DEC-2000; 2000WO-US42661.

XX 08-DEC-1999; 99US-0457498.

XX (FLUO-) FLUORO PROBE INC.

XX Luiken GA;

PI WPI; 2001-465198/50.

DR Treatment of pain associated with an interior disease site, involves

PT administering a pain-relieving target construct to the patient -

XX Claim 32; Page 20; 31pp; English.

XX The sequence represents the amino acid sequence of linker moiety #1 used

CC to make a pain-relieving target construct. The construct comprises a

CC pain-relieving agent linked to a ligand moiety that selectively binds to

CC or is taken up by the tissue associated with the painful interior disease

CC site. The construct is used for treating pain associated with an interior

CC disease site. Since the pain-relieving agent is delivered by the ligand

CC to the disease site intractable pain situated in the interior of the

CC body such as caused by various tumours, such as breast cancer, lung

CC cancer, Hodgkin's and non-Hodgkin's lymphoma, and neuroblastoma, and

CC ischaemic and diseased tissues can be managed using a lower level of the

CC pain relieving agent than is required when the pain-relieving agent is

CC injected in the free state.

XX SQ Sequence 5 AA;

Query Match 100.0%; Score 5; DB 22; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
 |||||
 Db 1 GGGGS 5

RESULT 20
 AAU05173
 ID AAU05173 standard; peptide; 5 AA.
 XX
 AC AAU05173;
 XX
 DT 24-OCT-2001 (first entry)
 XX
 DE Pain-relieving target construct, linker moiety #2.
 XX
 KW Pain relief; tumour; therapeutic; breast cancer; lung cancer;
 KW Hodgkin's lymphoma; neuroblastoma; ischaemic.
 XX
 OS Synthetic.
 XX
 PN WO200147512-A2.
 XX
 PD 05-JUL-2001.
 XX
 PF 06-DEC-2000; 2000WO-US42661.
 XX
 PR 08-DEC-1999; 99US-0457498.
 XX
 PA (FLUO-) FLUORO PROBE INC.
 XX
 PI Luiken GA;
 XX
 DR WPI; 2001-465198/50.
 XX
 PT Treatment of pain associated with an interior disease site, involves
 PT administering a pain-relieving target construct to the patient -
 XX
 PS Claim 32; Page 20; 31pp; English.
 XX
 CC The sequence represents the amino acid sequence of linker moiety #2 used
 CC to make a pain-relieving target construct. The construct comprises a
 CC pain-relieving agent linked to a ligand moiety that selectively binds to
 CC or is taken up by the tissue associated with the painful interior disease
 CC site. The construct is used for treating pain associated with an interior
 CC disease site. Since the pain-relieving agent is delivered by the ligand
 CC to the disease site intractable pain situated in the interior of the
 CC body such as caused by various tumours, such as breast cancer, lung
 CC cancer, Hodgkin's and non-Hodgkin's lymphoma, and neuroblastoma, and
 CC ischaemic and diseased tissues can be managed using a lower level of the
 CC pain relieving agent than is required when the pain-relieving agent is
 CC injected in the free state.
 XX
 SQ Sequence 5 AA;

Query Match 100.0%; Score 5; DB 22; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
 |||||
 Db 1 GGGGS 5

RESULT 21
 AAG63011
 ID AAG63011 standard; peptide; 5 AA.
 XX
 AC AAG63011;
 XX

DT 01-OCT-2001 (first entry)
 XX
 DE Peptide linker used to link NNT-1 and sCNTFRalpha.
 XX
 KW Neurotrophin-1; NNT-1; ciliary neurotrophic factor receptor alpha-chain;
 KW CNTFRalpha; fusion protein; gp130 receptor; LIFRbeta receptor;
 KW neurodegenerative disorder; nervous tissue; skeletal muscle;
 KW muscle mass; amyotrophic lateral sclerosis; Parkinson's disease;
 KW Huntington's disease; Obesity; cancer; fertility; endometriosis;
 KW blastocyst transplantation; haematopoiesis; thrombocytopaenia; retinitis;
 KW retinitis pigmentosa.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 XX
 FT Misc-difference 1..5 /note= "this sequence is repeated 1-10 times,
 FT preferably 2-4 times"
 XX
 PN WO200155219-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 26-JAN-2001; 2001WO-FR00254.
 XX
 PR 27-JAN-2000; 2000FR-0001035.
 XX
 PR 12-OCT-2000; 2000FR-0013090.
 XX
 PA (FABR) FABRE MEDICAMENT SA PIERRE.
 XX
 PI Eleon G, Gauchat J;
 XX
 DR WPI; 2001-457721/49.
 XX
 PT Fusion protein, used to modulate the activity of the gp130/LIFRbeta
 PT receptor in the treatment of neurodegenerative disorders, obesity and
 PT cancer, comprises an NNT-1 and an sCNTFRalpha protein and a linking
 PT element -
 XX
 PS Claim 4; Page 54; 65pp; French.
 XX
 CC AAG63011-14 represent peptide linkers used to produce a fusion protein,
 CC comprising neurotrophin-1 (NNT-1) and a soluble ciliary neurotrophic
 CC factor receptor alpha-chain (sCNTFRalpha) protein. The fusion protein is
 CC used to prepare a medicament for the modulation of the activity of the
 CC gp130/LIFRbeta receptor in cells expressing the receptor, preferably
 CC or skeletal muscle cells. It is also used to make a medicament to
 CC or skeletal muscle cells. It is also used to make a medicament to
 CC facilitate the proliferation or inhibit the differentiation of cloned
 CC cells, preferably embryonic stem cell clones. The fusion protein is
 CC also used as a medication for preventing or treating neurodegenerative
 CC disorders, for the repair or regeneration of nervous tissue or skeletal
 CC muscle, to maintain muscle mass in paralysed subjects, to treat or
 CC prevent amyotrophic lateral sclerosis, Parkinson's and Huntington's
 CC disease, obesity and its related disorders, and cancer. The fusion
 CC protein may also be used to improve fertility, especially to prevent
 CC endometriosis and/or facilitate blastocyst transplantation, to improve
 CC haematopoiesis, especially in treating thrombocytopaenia, and for the
 CC treatment of retinitis, especially retinitis pigmentosa.
 XX
 SQ Sequence 5 AA;

Query Match 100.0%; Score 5; DB 22; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
 |||||
 Db 1 GGGGS 5

RESULT 22
 AAE06269

PN	WO200136447-A2.
XX	
PD	25-MAY-2001.
XX	
PX	17-OCT-2000; 2000WO-US41224.
PF	
XX	
PR	20-OCT-1999; 99US-0421971.
PP	(SALK) SALK INST BIOLOGICAL STUDIES.
XX	
PA	Gage FH, Suhr ST, Gil EB, Senut MC;
PI	WPI; 2001-355608/37.
XX	
DR	
XX	Novel chimeric protein useful for modulating exogenous gene expression
PT	in subjects, comprises two functional protein units, each containing
PT	dimerization domain of steroid/thyroid hormone nuclear receptor
PT	superfamily -
XX	
XX	Disclosure; Page 13; 60pp; English.
PS	
XX	The present invention relates to a chimeric protein which consists of at
CC	least two functional protein units where each protein unit comprises the
CC	dimerisation domain of a member of the steroid/thyroid hormone nuclear
CC	receptor superfamily. The protein units are linked by a linker peptide,
CC	and the two protein units form a functional entity. The invention
CC	includes a method for modulating the expression of an exogenous gene in
CC	an organism or cell containing the chimeric protein. The cell or organism
CC	also contains a DNA construct comprising the exogenous gene under the
CC	control of a response element with which the chimeric protein interacts.
CC	The response element controls the expression of the exogenous gene. The
CC	method also involves administering to the subject or cell an effective
CC	amount of an exogenous ligand for at least one functional unit of the
CC	chimeric protein. The chimeric protein is useful for modulating the
CC	expression of an exogenous gene in a subject organism. The present
CC	sequence represents a linker peptide which can be used in the
CC	construction of the chimeric protein of the invention.
XX	
XX	Sequence 5 AA;
SQ	
	Query Match 100.0%; Score 5; DB 22; Length 5;
	Best Local Similarity 100.0%; Pred. No. 9.3e+05;
	Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 GGGGS 5
DB	1 GGGGS 5
RESULT 24	
AAG62613	ID AAG62613 standard; peptide; 5 AA.
XX	
AC	AAG62613;
DT	06-SEP-2001 (first entry)
XX	
DE	Linker peptide.
XX	
KW	Linker; capillary array-based sample screening; bioactive molecule;
KM	enzyme; antibacterial drug.
XX	
OS	Synthetic.
XX	
PN	WO200138583-A2.
XX	
PD	31-MAY-2001.
XX	
XX	
PF	22-NOV-2000; 2000WO-US32208.
XX	
PR	22-NOV-1999; 99US-0444112.
PR	12-OCT-2000; 2000US-0687219.
XX	

PA (DIVE-) DIVERSA CORP.
 XX Lafferty WM, Short JM, Keller M;
 XX WPI; 2001-367710/38.
 DR Identifying a bioactivity or biomolecule of interest, involves
 PT culturing a substrate labeled with a detectable molecule and a
 PT recombinant clone in a capillary tube of capillary array, and detecting
 PT the signal -
 XX Disclosure; Page 79; 133pp; English.
 XX The present invention relates to a method of identifying a bioactivity or
 CC biomolecule of interest, involving introducing a labeled substrate and a
 CC recombinant clone into a capillary tube of a capillary array, culturing
 CC the tube to allow the substrate and clone to interact and provide a
 CC detectable signal and detecting the signal to identify the capillaries
 CC containing the activity or molecule of interest. This can be used to
 CC identify enzymes for use in the biotechnology and chemical industries,
 CC and in drug discovery, particularly of antibacterial compounds. The
 CC present sequence is a linker peptide described in the exemplification of
 CC the invention.
 XX Sequence 5 AA;
 SQ Query Match 100.0%; Score 5; DB 22; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGGS 5
 Db |||||
 1 GGGGS 5
 RESULT 25
 AAB81033
 ID AAB81033 standard; peptide; 5 AA.
 AC AAB81033;
 XX 15-JUN-2001 (first entry)
 DT Linker peptide used in the production of a PQQGDH fusion protein.
 DE Pyroloquinolinequinone glucose dehydrogenase; PQQGDH; fusion protein;
 XX linker; glucose.
 OS Synthetic.
 XX JP2001037483-A.
 PN 13-FEB-2001.
 PD 30-JUL-1999; 99JP-0216459.
 XX 30-JUL-1999; 99JP-0216459.
 PR (HAVA/) HAYADE H.
 PA WPI; 2001-285600/30.
 DR Fused proteins having two pyrroloquinolinequinone glucose
 PT dehydrogenases, used for quantitative determination of glucose in
 PT samples -
 XX Claim 5; Page 2; 7pp; Japanese.
 XX This invention relates to a fusion protein consisting of two water
 CC soluble pyrroloquinolinequinone glucose dehydrogenase (PQQGDH) subunits
 CC derived from *Acinetobacter calcoaceticus*, bound via a linker peptide. The
 CC fusion protein is used for quantitative determination of glucose in
 CC clinical samples and foods. The present sequence represents a linker

CC peptide which forms part of the fusion protein of the invention.
 XX Sequence 5 AA;
 SQ Query Match 100.0%; Score 5; DB 22; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGGS 5
 Db |||||
 1 GGGGS 5
 RESULT 26
 AAB74580
 ID AAB74580 standard; peptide; 5 AA.
 XX AAB74580;
 AC 05-JUN-2001 (first entry)
 DT Context-dependent functional entity spacer peptide #2.
 XX Context-dependent functional entity; efficacy enhancement; inflammation;
 KW selective recognition domain; thrombogenic potential; cancer; infection;
 KW arthritis; atherosclerosis; burn.
 OS Synthetic.
 XX WO200125413-A1.
 PN 12-APR-2001.
 PD 04-OCT-2000; 2000WO-US27794.
 PF 04-OCT-1999; 99US-0411067.
 PR (NUVA-) NUVAS LLC.
 PA Houston LL;
 PI WPI; 2001-258295/26.
 DR Enhancing the efficacy of active and diagnostic agents and treating
 XX pathological conditions in a subject by administering the agent with a
 PT context-dependent functional entity, useful for treating e.g. malignant
 PT neoplasms -
 XX Disclosure; Page 23; 49pp; English.
 PS The present invention describes a method of enhancing the efficacy of an
 XX agent by coadministering the agent and a context-dependent functional
 CC entity. The latter comprises a substructure with thrombogenic potential
 CC and a selective recognition domain. This is useful in the treatment of
 CC cancer, arthritis, infections, neovascular glaucoma, inflammatory
 CC conditions, dermatitis, endometriosis, atherosclerosis, vascular
 CC restenosis and conditions associated with granulation tissues, such as
 CC burns and pyogenic granuloma. The present sequence is an example of a
 CC spacer peptide for use in the invention.
 XX Sequence 5 AA;
 SQ Query Match 100.0%; Score 5; DB 22; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGGS 5
 Db |||||
 1 GGGGS 5
 RESULT 27
 AAB68571

ID AAB68571 standard; peptide; 5 AA.
 XX AC AAB68571;
 XX 20-APR-2001 (first entry)
 XX DT
 XX Linker peptide #1.
 XX DR
 XX Linker peptide #1.
 XX XX Transgenic plant; expression cassette.
 XX XX Unidentified.
 XX XX WO200105936-A2.
 XX XX
 XX XX 25-JAN-2001.
 XX XX
 XX XX 18-JUL-2000; 2000WO-US19721.
 XX XX
 XX XX 19-JUL-1999; 99US-0144513.
 XX XX 07-JUL-2000; 2000US-0611736.
 XX XX (INNE-) INNES CENT JOHN.
 XX XX Christou P, Kohli A;
 XX XX WPI; 2001-147332/15.
 XX XX
 XX XX Producing population of transgenic plants involves transforming plant
 XX XX cell with minimal transgene expression cassette by direct DNA transfer
 XX XX techniques and regenerating transgenic plants from transformed cells -
 XX XX
 XX XX Disclosure; Page 11; 42pp; English.
 XX XX
 XX XX The present invention relates to a method for producing a population of
 XX XX transgenic plants, comprising transforming intact plant cells with
 XX XX minimal transgene expression cassettes by direct DNA transfer techniques,
 XX XX and regenerating a population of transgenic plants from transformed plant
 XX XX cells. The present sequence is a linker peptide, which can be used in the
 XX XX expression cassette of the present invention.
 XX XX
 XX XX Sequence 5 AA;
 XX XX
 XX XX Query Match 100.0%; Score 5; DB 22; Length 5;
 XX XX Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 XX XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX XX
 XX QY 1 GGGGS 5
 XX DB |||||
 XX 1 GGGGS 5
 XX
 XX RESULT 28
 XX AAB50848
 XX ID AAB50848 standard; Peptide; 5 AA.
 XX XX
 XX XX AAB50848;
 XX XX
 XX XX 14-MAR-2001 (first entry)
 XX XX
 XX XX Fluorescent protein indicator linker moiety #1.
 XX DE
 XX XX Fluorescent protein indicator; green fluorescent protein; GFP;
 XX XX linker moiety; sensor; calmodulin-binding domain.
 XX KW
 XX XX Synthetic.
 XX OS
 XX XX WO2000071565-A2.
 XX XX
 XX XX 30-NOV-2000.
 XX XX
 XX XX 17-MAY-2000; 2000WO-US13684.
 XX PF
 XX XX 21-MAY-1999; 99US-0316919.
 XX PI

PR 21-MAY-1999; 99US-0316920.
 XX
 XX (REGC) UNIV CALIFORNIA.
 XX PI
 XX Tsien RY, Baird GA;
 XX DR
 XX WPI; 2001-032017/04.
 XX
 XX Novel fluorescent proteins comprising a sensor protein inserted into
 XX them, useful for measuring the response of a sensor biological,
 XX chemical, electrical or physiological parameter in vivo or in vitro -
 XX Disclosure; Page 34; 94pp; English.
 XX
 XX The present sequence is a linker moiety used in the construction of
 XX a fluorescent protein indicator. The indicator comprises a sensor
 XX polypeptide that is responsive to a chemical, biological, electrical
 XX or physiological parameter, and a fluorescent protein functional group.
 XX The sensor polypeptide is operatively inserted into the fluorescent
 XX moiety. The fluorescent indicator is useful for detecting the presence of
 XX a response inducing member in a sample. The method involves contacting
 XX the sample with the indicator and detecting a change in fluorescence, in
 XX which a change is indicative of the effect of the parameter on the sensor
 XX polypeptide. The novel fluorescent proteins are advantageous due to their
 XX reduced size as compared to the FRET (fluorescence resonance energy
 XX transfer)-based sensors.
 XX
 XX Sequence 5 AA;
 XX
 XX Query Match 100.0%; Score 5; DB 22; Length 5;
 XX Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 1 GGGGS 5
 XX DB |||||
 XX 1 GGGGS 5
 XX
 XX RESULT 29
 XX AAB55527
 XX ID AAB55527 standard; Peptide; 5 AA.
 XX XX
 XX XX AAB55527;
 XX DT
 XX 07-MAR-2001 (first entry)
 XX
 XX Peptide linker #1.
 XX DE
 XX XX Hematopoietic cell; PDZ; PL; autoimmune disease; inflammation;
 XX KW allergy; asthma; multiple sclerosis; cancer; infection.
 XX OS
 XX Unidentified.
 XX XX
 XX XX WO2000069896-A2.
 XX XX
 XX XX 23-NOV-2000.
 XX XX
 XX XX 12-MAY-2000; 2000WO-US13161.
 XX XX
 XX XX 14-MAY-1999; 99US-0134114.
 XX PR 14-MAY-1999; 99US-0134117.
 XX PR 14-MAY-1999; 99US-0134118.
 XX PR 21-OCT-1999; 99US-0160860.
 XX PR 29-OCT-1999; 99US-0162498.
 XX PR 13-DEC-1999; 99US-0170453.
 XX PR 14-JAN-2000; 2000US-0176195.
 XX PR 14-FEB-2000; 2000US-0182296.
 XX PR 11-APR-2000; 2000US-0196460.
 XX PR 11-APR-2000; 2000US-0196527.
 XX XX
 XX XX (ARBO-) ARBOR VITA CORP.
 XX PF
 XX XX Lu PS;

XX WPI; 2001-080245/09.
 XX
 XX Modulating a biological function of an endothelial cell or
 PT hematopoietic cell, useful for treating autoimmune diseases and
 PT infectious diseases, by administering an antagonist that inhibits
 PT binding between a PDZ protein and a PL protein -
 XX
 XX Disclosure; Page 106; 141pp; English.
 XX
 XX The present invention relates to a new method for modulating a
 CC biological function of an endothelial cell or hematopoietic cell. The
 CC method involves introducing into a cell, an antagonist that inhibits
 CC binding between a PDZ protein and a PL protein. The inhibitor is used
 CC to treat a disease mediated by hematopoietic cells, e.g. autoimmune
 CC disease. It may also be used to prevent transplantation rejection of
 CC a solid organ transplant. The method may also be used in the treatment
 CC of inflammation, allergy, inflammatory bowel diseases, ulcerative
 CC colitis, ileitis, psoriasis, asthma, atopic dermatitis, autoimmune
 CC diseases (e.g. rheumatoid arthritis, multiple sclerosis,
 CC insulin-dependent diabetes, Hashimoto thyroiditis, osteoarthritis,
 CC graft rejection, transplantation rejection), atherosclerosis, cancers,
 CC infectious diseases, ischemia, vasculitis and Crohn's disease.
 XX
 XX Sequence 5 AA;
 SQ
 Query Match 100.0%; Score 5; DB 22; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGGS 5
 DB |||||
 1 GGGGS 5

RESULT 30
 AAB57374
 ID AAB57374 standard; Peptide; 5 AA.
 XX
 AC AAB57374;
 XX
 DT 12-MAR-2001 (first entry)
 DE Flexible polylinker #1.
 XX
 KW Endothelial cell; hematopoietic cell; PDZ domain protein;
 KW PL domain protein; leukocyte activation; synapse formation;
 KW transmembrane neurotransmitter receptor; autoimmune disease;
 KW transplantation rejection; inflammation; allergy;
 KW inflammatory bowel disease; ulcerative colitis; ileitis; psoriasis;
 KW asthma; atopic dermatitis; atherosclerosis; cancer; infectious disease;
 KW ischaemia; vasculitis; Crohn's disease.
 XX
 OS Unidentified.
 XX
 PN WO200069897-A2.
 PD
 XX 23-NOV-2000.
 XX
 PF 12-MAY-2000; 2000WC-US13166.
 XX
 PR 14-MAY-1999; 99US-0134114.
 PR 14-MAY-1999; 99US-0134117.
 PR 14-MAY-1999; 99US-0134118.
 PR 21-OCT-1999; 99US-0160860.
 PR 29-OCT-1999; 99US-0162498.
 PR 13-DEC-1999; 99US-0170453.
 PR 14-JAN-2000; 2000US-0176195.
 PR 14-FEB-2000; 2000US-0182296.
 PR 11-APR-2000; 2000US-0196460.
 PR 11-APR-2000; 2000US-0196527.
 XX
 PA (ARBO-) ARBOR VITA CORP.

XX Lu PS;
 PI
 XX WPI; 2001-025003/03.
 DR
 XX New inhibitors of binding of a PDZ protein and PL protein for
 PT inhibiting T cell-mediated response by hematopoietic cells, or for
 PT treating diseases characterized by inflammatory and humoral immune
 PT responses, e.g. inflammation, cancer -
 XX
 XX Disclosure; Page 86; 139pp; English.
 XX
 XX The present invention relates to a method for modulating a biological
 CC function of an endothelial cell or hematopoietic cell, comprises
 CC introducing into a cell an antagonist that inhibits binding between a
 CC PDZ domain protein and a PL domain protein to result in inhibition of
 CC leukocyte activation. PDZ domains of proteins are named after three
 CC prototypical proteins: PSD95, Drosophila large disc protein and Zonula
 CC Occludin 1 protein. PDZ domain proteins are involved in synapse formation
 CC by organising transmembrane neurotransmitter receptors through
 CC intracellular interactions. The inhibitors identified by the present
 CC invention can be used to treat a disease mediated by hematopoietic
 CC cells, e.g. autoimmune disease, inflammation, allergy (e.g. drug
 CC allergies), inflammatory bowel diseases, ulcerative colitis, ileitis,
 CC psoriasis, respiratory allergic diseases (e.g. asthma), atopic
 CC dermatitis, autoimmune diseases (e.g. rheumatoid arthritis, multiple
 CC sclerosis, insulin-dependent diabetes, cancers, infectious diseases (e.g.
 CC osteoarthritis), atherosclerosis, vasculitis and Crohn's disease. The
 CC viral infection), ischaemia, vasculitis and Crohn's disease. The
 CC inhibitors can also be used to prevent transplantation rejection of a
 CC solid organ transplant. The present sequence is a flexible polylinker
 CC used in the construction of fusion polypeptides of the PDZ-PL
 CC antagonists.
 XX
 XX Sequence 5 AA;
 SQ
 Query Match 100.0%; Score 5; DB 22; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGGS 5
 DB |||||
 1 GGGGS 5

RESULT 31
 AAB57784
 ID AAB57784 standard; Peptide; 5 AA.
 XX
 AC AAB57784;
 XX
 DT 12-MAR-2001 (first entry)
 DE Flexible polylinker #1.
 XX
 KW Endothelial cell; hematopoietic cell; PDZ domain protein;
 KW PL domain protein; leukocyte activation; synapse formation;
 KW transmembrane neurotransmitter receptor; autoimmune disease;
 KW transplantation rejection; inflammation; allergy;
 KW inflammatory bowel disease; ulcerative colitis; ileitis; psoriasis;
 KW asthma; atopic dermatitis; atherosclerosis; cancer; infectious disease;
 KW ischaemia; vasculitis; Crohn's disease.
 XX
 OS Unidentified.
 XX
 PN WO200069898-A2.
 PD
 XX 23-NOV-2000.
 XX
 PF 12-MAY-2000; 2000WO-US13205.
 XX
 PR 14-MAY-1999; 99US-0134114.
 PR 14-MAY-1999; 99US-0134117.

14-MAY-1999; 99US-0134118.
 21-OCT-1999; 99US-0160860.
 29-OCT-1999; 99US-0162498.
 13-DEC-1999; 99US-0170453.
 14-JAN-2000; 2000US-0176195.
 14-FEB-2000; 2000US-0182296.
 11-APR-2000; 2000US-0196460.
 11-APR-2000; 2000US-0196527.
 (ARBO-) ARBOR VITA CORP.
 Lu PS;
 WPI; 2001-061214/07.
 Modulating a biological function of a hematopoietic cell for treating an allergic response, or diseases mediated by immune system cells, comprises introducing into the cell a PDZ-PL interaction enhancer or inhibitor -
 Disclosure; Page 107; 143pp; English.
 The present invention relates to a method for modulating a biological function of an endothelial cell or hematopoietic cell, comprises introducing into a cell an antagonist that inhibits binding between a PDZ domain protein and a PL domain protein to result in inhibition of leukocyte activation. PDZ domains of proteins are named after three prototypal proteins: PS95, Drosophila large disc protein and Zonula Occludin 1 protein. PDZ domain proteins are involved in synapse formation by organising transmembrane neurotransmitter receptors through intracellular interactions. The inhibitors identified by the present invention can be used to treat a disease mediated by hematopoietic cells, e.g. autoimmune disease, inflammation, allergy (e.g. drug allergies), inflammatory bowel diseases, ulcerative colitis, ileitis, psoriasis, respiratory allergic diseases (e.g. asthma), atopic dermatitis, autoimmune diseases (e.g. rheumatoid arthritis, multiple sclerosis, insulin-dependent diabetes, infectious diseases (e.g. osteoarthritis), atherosclerosis, cancers, thyroiditis, viral infection), ischaemia, vasculitis and Crohn's disease. The inhibitors can also be used to prevent transplantation rejection of a solid organ transplant. The present sequence is a flexible polypeptide used in the construction of fusion polypeptides of the PDZ-PL antagonists.

Query Match 100.0%; Score 5; DB 22; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DY 1 GGGGS 5
 |||||
 Db 1 GGGGS 5

RESULT 32
 ABP56192
 ID ABP56192 standard; Peptide; 5 AA.
 XX AC AAE28934;
 XX DT 27-JAN-2003 (first entry)
 XX DE Linker peptide used in the invention.
 XX KW Baculovirus; protein delivery; peptide display; capsid; linker.
 XX OS Unidentified.
 XX PN WO200272814-A2.
 XX PD 19-SEP-2002.
 XX PF 12-MAR-2002; 2002WO-GB01115.
 XX PR 12-MAR-2001; 2001GB-0006063.
 XX PS Synthetic.

PN WO200261105-A2.
 XX 08-AUG-2002.
 PD 01-FEB-2002; 2002WO-EP01633.
 PF 02-FEB-2001; 2001US-265594P.
 PR (INSP) INST PASTEUR.
 XX (CNRS) CENT NAT RECH SCI.
 PA Edelman L, Jacotot E, Briand J;
 XX WPI; 2002-619260/66.
 DR
 XX New chimeric bifunctional molecules that target specific cells and regulate the apoptosis function of the permeability transition pore complex of the mitochondria, useful for treating or preventing e.g. cancer or ischemia -
 Example 2; Page 31; 76pp; English.
 The present invention describes a chimeric bifunctional molecule (I) comprising at least a first functional molecule covalently linked to a second functional molecule, which is able to modulate the activity of the permeability transition pore complex (PTPC) of the mitochondria. (I) has the function of specifically targeting and entering a tissue cell population. The second functional molecule has the function of specifically targeting, and inducing or preventing the death of the cells by apoptosis by regulating the opening or the closing of the PTPC of the mitochondria or its fragment. (I) has virucide, neuroprotective, vasotropic and cytostatic activities, and can be used as a mitochondrial permeability transition pore complex (PTPC) modulator. (I) is useful for treating or preventing a pathological infection or disease. (I) is also useful for regulating cell death regulatory molecules, specifically the apoptogenic function of the PTPC, for treating e.g. cancer, ischaemia, neurodegenerative diseases, fulminant hepatitis or viral infections. The present sequence represents a linker peptide which is used in an example from the present invention.

Query Match 100.0%; Score 5; DB 23; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DY 1 GGGGS 5
 |||||
 Db 1 GGGGS 5

RESULT 33
 AAE28934
 ID AAE28934 standard; peptide; 5 AA.
 XX AC AAE28934;
 XX DT 27-JAN-2003 (first entry)
 XX DE Linker peptide used in the invention.
 XX KW Baculovirus; protein delivery; peptide display; capsid; linker.
 XX OS Unidentified.
 XX PN WO200272814-A2.
 XX PD 19-SEP-2002.
 XX PF 12-MAR-2002; 2002WO-GB01115.
 XX PR 12-MAR-2001; 2001GB-0006063.
 XX PS

PA (ARKT-) ARK THERAPEUTICS LTD.

PI Yia-Herttua S, Airenne KJ;

XX WPI; 2002-723349/78.

XX New engineered baculoviruses, useful for the delivery of proteins into
PT the nucleus, libraries or peptide displays, e.g. in selecting target
PT genes, comprises capsids modified to display heterologous proteins -

XX Disclosure; Page 6; 17pp; English.

XX The invention relates to a baculovirus, whose capsid has been modified
CC to display one or more heterologous peptides or proteins. The baculovirus
CC is useful for the delivery of a peptide or protein into the nucleus. It
CC is also useful in libraries (genomic or cDNA) and peptide displays,
CC particularly in selecting target genes, in the functional analysis of
CC genes and their products, or screening these libraries. The present
CC sequence is a linker peptide used in the invention.

XX Sequence 5 AA;

Query Match 100.0%; Score 5; DB 23; Length 5;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5

|||||

Db 1 GGGGS 5

RESULT 34

AAE28173

ID AAE28173 standard; peptide: 5 AA.

AC AAE28173;

DT 27-DEC-2002 (first entry)

XX Linker peptide #5 used to link zinc finger proteins.

XX Phenotype; zinc finger protein; ZFP; gene expression; tumour progression;
KW transformation; inflammatory response; neurological disorder; linker.

XX Synthetic.

XX US2002094529-A1.

XX 18-JUL-2002.

XX 28-AUG-2001; 2001US-0941450.

XX 14-SEP-1999; 99US-0395448.

XX (CASE/) CASE C C.

XX (URNO/) URNOV F.

XX Case CC, Urnov F;

XX WPI; 2002-673639/72.

XX Identifying a gene, useful for exploring biological processes, e.g.
PT transformation, tumor progression, or inflammatory response, comprises
PT contacting a cell comprising a putative gene sequence with a zinc
PT finger protein -

XX Disclosure; Page 35; 46pp; English.

XX The present invention relates to methods and compositions for identifying
CC a gene. The method involves contacting a cell comprising a putative gene
CC sequence [PGS] with an exogenous molecule that binds to and modulates the
CC expression of the PGS, and assaying the cell for at least one selected
CC phenotype, where if one or more of the selected phenotypes are observed,

CC the PGS is identified as a gene. The methods include the use of exogenous
CC molecules such as zinc finger proteins (ZFPs) which are capable of
CC binding to and modulating expression of gene transcription, targeted to
CC PGS, followed by assay of one or more selected phenotypes. The ZFPs are
CC useful for determining the phenotypic consequences and function of gene
CC expression. The method is useful to explore biological processes such as
CC transformation, tumor progression, inflammatory response, neurological
CC disorders, etc. The method can very rapidly identify genes that are up or
CC down-regulated in different tissues or in response to specific stimuli.
CC The present sequence is a linker peptide used to link two zinc finger
CC proteins. This peptide is used to illustrate the method of the invention.

XX Sequence 5 AA;

Query Match 100.0%; Score 5; DB 23; Length 5;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5

|||||

Db 1 GGGGS 5

RESULT 35

ABB79970

ID ABB79970 standard; Peptide: 5 AA.

XX ABB79970;

DT 19-DEC-2002 (first entry)

XX Linker peptide useful in fluorescent targeting construct.

XX Tumour; cancer; infection; heart attack; stroke; necrosis;
KW ischaemia; diagnosis; fluorescence; linker.

XX Synthetic.

XX US2001055566-A1.

XX 27-DEC-2001.

XX 09-APR-2001; 2001US-0832297.

XX 15-OCT-1998; 98US-0173190.

XX 28-JUL-1999; 99US-0362805.

XX (LUIK/) LUIKEN G A.

XX Luiken GA;

XX WPI; 2002-403040/43.

XX In vivo diagnosis of diseased tissue, useful e.g. for defining extent
PT of tumours during surgical resection, based on fluorescence from a
PT specifically bound targeting construct -

XX Disclosure; Page 10; 13pp; English.

XX The present sequence is that of a flexible spacer peptide linker
CC that can be used to link a fluorescent moiety to a ligand moiety in
CC fluorescent targeting constructs of the invention. A claimed
CC method for in vivo diagnosis of diseased tissue involves:
CC irradiating a body part with light having at least one excitation
CC wavelength of 401-500 nm; directly viewing fluorescence emitted in
CC response to light from a fluorescent targeting construct
CC administered to the subject and which has specifically bound to
CC and/or been taken up by the diseased tissue; and determining the
CC location and/or surface area of the diseased tissue from the
CC fluorescence provided by the targeting construct. The ligand
CC moiety is especially a tumour-avid moiety. The method is
CC especially used to diagnose any of a wide range of tumours
CC (melanoma, insulinoma, pancreatic tumours, small cell and non-small

C cell lung cancer, lymphoma, or ovary, pituitary, pancreas or adrenal
 C cancer, brain tumour, colorectal cancer, cutaneous melanoma,
 C epithelial cancer, lung carcinoma, testicular germ cell tumour, or
 C breast cancer), especially during surgery to define the outline and
 C size of the tissue that needs to be removed. It can also be used to
 C detect infections (bacterial, viral or fungal), pre-cancerous states,
 C heart attack, stroke and necrotic or ischaemic conditions (all
 C claimed).

Q Sequence 5 AA;
 Query Match 100.0%; Score 5; DB 23; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
 |||||
 Db 1 GGGGS 5

RESULT 36
 ABB79971
 ID ABB79971 standard; Peptide; 5 AA.

AC ABB79971;
 XX 19-DEC-2002 (first entry)
 DT Linker peptide useful in fluorescent targeting construct.
 DE
 XX Tumour; cancer; infection; heart attack; stroke; necrosis;
 KW ischaemia; diagnosis; fluorescence; linker.
 XX Synthetic.

Key Location/Qualifiers
 Misc-difference 1..5 /note= "the sequence can be repeated n times"

US2001055566-A1.

27-DEC-2001.

09-APR-2001; 2001US-0832297.

15-OCT-1998; 98US-0173190.

28-JUL-1999; 99US-0362805.

(LUIK/) LUIKEN G A.

Luiken GA;

WPI; 2002-403040/43.

In vivo diagnosis of diseased tissue, useful e.g. for defining extent
 of tumors during surgical resection, based on fluorescence from a
 specifically bound targeting construct -

Disclosure; Page 10; 13pp; English.

The present sequence is that of a flexible spacer peptide linker
 that can be used to link a fluorescent moiety to a ligand moiety in
 fluorescent targeting constructs of the invention. A claimed

method for in vivo diagnosis of diseased tissue involves:
 irradiating a body part with light having at least one excitation
 wavelength of 401-500 nm; directly viewing fluorescence emitted in
 response to light from a fluorescent targeting construct
 administered to the subject and which has specifically bound to
 and/or been taken up by the diseased tissue; and determining the
 location and/or surface area of the diseased tissue from the
 fluorescence provided by the targeting construct. The ligand
 moiety is especially a tumour-avid moiety. The method is
 especially used to diagnose any of a wide range of tumours

CC (melanoma, insulinoma, pancreatic tumours, small cell and non-small
 CC cell lung cancer, lymphoma, or ovary, pituitary, pancreas or adrenal
 CC cancer, brain tumour, colorectal cancer, cutaneous melanoma,
 CC epithelial cancer, lung carcinoma, testicular germ cell tumour, or
 CC breast cancer), especially during surgery to define the outline and
 CC size of the tissue that needs to be removed. It can also be used to
 CC detect infections (bacterial, viral or fungal), pre-cancerous states,
 CC heart attack, stroke and necrotic or ischaemic conditions (all
 CC claimed).

QY Sequence 5 AA;
 Query Match 100.0%; Score 5; DB 23; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
 |||||
 Db 1 GGGGS 5

RESULT 37
 ABB80672
 ID ABB80672 standard; Peptide; 5 AA.

AC ABB80672;
 XX 29-NOV-2002 (first entry)
 DT Synthetic linker peptide.
 DE
 XX Molecular antigen array; vaccine; antigen; antimicrobial;
 KW molecular scaffold; amyloid beta; Abeta 1-42; influenza;
 KW graft versus host disease; Ige-mediated allergic reaction; anaphylaxis;
 KW adult respiratory distress syndrome; ARDS; Crohn's disease;
 KW allergic asthma; acute lymphoblastic leukaemia; non-Hodgkin's lymphoma;
 KW Grave's disease; systemic lupus erythematosus; osteoporosis;
 KW inflammatory immune disease; myasthenia gravis; multiple sclerosis;
 KW immunoproliferative disease lymphadenopathy; Alzheimer's disease;
 KW angioimmunoproliferative lymphadenopathy; immunoblastic lymphadenopathy;
 KW rheumatoid arthritis; diabetes; infectious disease; factor Xa;
 KW enterokinase; cysteine-containing linker.

Synthetic.

WO200256907-A2.

25-JUL-2002.

21-JAN-2002; 2002WO-IB00168.

19-JAN-2001; 2001US-262379P.

04-MAY-2001; 2001US-288549P.

05-OCT-2001; 2001US-326998P.

07-NOV-2001; 2001US-331045P.

(CYTO-) CYTOS BIOTECHNOLOGY AG.

(NOVS) NOVARTIS PHARMA AG.

(MAUR/) MAURER P.

(LECH/) LECHNER F.

(ORTM/) ORTMANN R.

(LUEO/) LUEOEND R.

(STAU/) STAUFENBIEL M.

(FREY/) FREY P.

Maurer P, Lechner F, Ortman R, Lueoend R, Staufenbiel M, Frey P;

Renner WA, Bachmann M, Tissot A, Sebbel P, Piossek C;

WPI; 2002-636514/68.

Molecular antigen array used in the production of vaccines for
 infectious diseases -

PS Disclosure; Page 49; 418pp; English.

XX The invention relates to a composition comprising: (a) a non-natural molecular scaffold comprising: (i) a core particle selected from: (1) a core particle of a non-natural origin; and (2) a core particle of natural origin; and (ii) an organiser comprising at least one first attachment site, where the organiser is connected to the core particle by at least one covalent bond; (b) an antigen or antigenic determinant with at least one second attachment site, where the antigen or antigenic determinant is amyloid beta peptide (Abeta 1-42) or its fragment, and where the second attachment site is selected from: (i) an attachment site not naturally occurring with the antigen or antigenic determinant; and (ii) an attachment site naturally occurring with the antigen or antigenic determinant, where the second attachment site is capable of association through at least one non-peptide bond to the first attachment site; and where the antigen or antigenic determinant and the scaffold interact through the association to form an ordered and repetitive antigen array. Also included is a process for producing a non-naturally occurring ordered and repetitive antigen array. The composition is used in immunisation and as a vaccine for diseases such as influenza, graft versus host disease, IGE-mediated allergic reactions, allergic adult respiratory distress syndrome (ARDS), Crohn's disease, allergic asthma, acute lymphoblastic leukaemia, non-Hodgkin's lymphoma, Grave's disease, systemic lupus erythematosus, inflammatory immune diseases, myasthenia gravis, immunoproliferative disease lymphadenopathy, angioimmunoproliferative lymphadenopathy, immunoblastic lymphadenopathy, rheumatoid arthritis, diabetes, multiple sclerosis, Alzheimer's disease, osteoporosis and infectious diseases. The antigens are modified to possess a cleavage site (enterokinase or factor Xa) and a Cysteine-containing N- or C-terminal linker peptide which serves as the attachment point to a virus like particle or bacterial protein (the scaffold protein). The present sequence is a cysteine-containing linker peptide used in the molecular antigen array.

SQ Sequence 5 AA;

Query Match 100.0%; Score 5; DB 23; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|||

Db 1 GGGGS 5

RESULT 38
AAE25959
ID AAE25959 standard; peptide; 5 AA.
XX AC AAE25959;
XX 15-NOV-2002 (first entry)
DT Linker peptide #3 used to produce single chain (scFv) antibodies.
DE Immunoglobulin; kinase insert domain-containing receptor; KDR;
XX foetal liver kinase; Fk-1 receptor; vascular endothelial growth factor;
KW VEGF; tumour growth; heavy chain variable region; VH; angiogenesis;
KW p1C11; scFv antibody.
XX Unidentified.
OS US2002064528-A1.
XX 30-MAY-2002.
PN 12-OCT-2001; 2001US-0976787.
XX 28-JAN-2000; 2000US-0493539.
PR (ZHUZ//) ZHU Z.
PA (WITT//) WITTE L.
XX

PI Zhu Z, Witte L;
XX WPI: 2002-599175/63.
DR N-PSDB; AAD42823.
XX Novel immunoglobulin molecule for reducing tumor growth, binds to kinase insert domain-containing receptor with an affinity comparable to human vascular endothelial growth factor, and neutralizes activation of KDR -
XX Disclosure; Page 14; 34pp; English.
PS The present invention relates to novel immunoglobulin molecules that bind to kinase insert domain-containing receptor (KDR) (a human homologue of mouse foetal liver kinase (FLK)-1 receptor) with an affinity comparable to human vascular endothelial growth factor (VEGF) and that neutralises activation of KDR. Sequences of the invention are useful for neutralising the activation of KDR, for reducing tumour growth and for inhibiting angiogenesis. The present sequence is a linker peptide used to produce single chain (scFv) antibodies.
XX Sequence 5 AA;
SQ Query Match 100.0%; Score 5; DB 23; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|||

Db 1 GGGGS 5

RESULT 39
ABJ05006
ID ABJ05006 standard; Peptide; 5 AA.
XX AC ABJ05006;
XX 07-NOV-2002 (first entry)
DT VH and VL antibody peptide linker SEQ ID No 1.
DE Immunosuppressive; antiinflammatory; affinity; Kd; binding; PDZ domain; ligand; Ki; inhibitor; K-enhancer; leukocyte; autoimmune disease; inflammatory; humoral immune response; inflammation; peptide linker.
XX Unidentified.
OS WO200231512-A2.
XX 18-APR-2002.
PN 11-OCT-2001; 2001WO-US32150.
XX 13-OCT-2000; 2000US-0688017.
PR (ARBO-) ARBOR VITA CORP.
XX Rabinowitz JD, Lu PS, Schweizer J;
XX WPI: 2002-416878/44.
XX Assays for determining the affinity of binding between a PDZ domain and a ligand, and determining the Ki of an inhibitor of the binding, comprises using a polypeptide comprising a PDZ domain and a non-PDZ domain -
XX Disclosure; Page 129; 164pp; English.
PS The invention relates to methods and reagents for determining the apparent affinity (Kd) of binding between a PDZ domain and a ligand. The invention also relates to methods and reagents for determining the Ki of an inhibitor of binding between a PDZ domain and a ligand, identifying an

agent that enhances binding of a PDZ domain and a ligand, and determining the potency (K-enhancer) of binding between a PDZ domain and a ligand, by determining the ligand bound with an immobilised polypeptide comprising a PDZ domain and a non-PDZ domain on a surface. The modulator (preferably, an inhibitor) of interaction between PDZ and PL is useful for treating a disease characterised by leukocyte activation, e.g., an autoimmune disease that is characterised by inflammatory or humoral immune response, and for reducing inflammation in a subject. This sequence represents a peptide linker of a VH and VL antibody relating to the PDZ domain CC peptides and ligands of the invention.

Sequence 5 AA;
Query Match 100.0%; Score 5; DB 23; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2y 1 GGGS 5
|||||

3b 1 GGGS 5
|||||

RESULT 40

ABG31035
ID ABG31035 standard; Peptide; 5 AA.

XX AC
XX AC
XX AC

05-NOV-2002 (first entry)

Binding domain-immunoglobulin fusion protein associated linker #1.

Immunosuppressive; antirheumatic; antithyroid; antidiabetic; mouse; neuroprotective; gene therapy; single chain antibody; variable fragment; scFv; binding domain-immunoglobulin fusion protein; B-cell disorder; malignant condition; rheumatoid arthritis; myasthenia gravis; psoriasis; Grave's disease; Hashimoto's thyroiditis; type I diabetes mellitus; multiple sclerosis; systemic lupus erythematosus; Sjogrens syndrome; immune thrombocytopenic purpura; scleroderma; cancer; Chron's disease; ulcerative colitis; inflammatory bowel disease; immunological effector; cell mediated cytotoxicity; complement dependent cytotoxicity; complement fixation.

Synthetic.

WO200255910-A1.

25-JUL-2002.

17-JAN-2002; 2002WO-US01487.

17-JAN-2001; 2001US-0765208.

(GENE-) GENE-CRAFT INC.

Ledbetter JA, Hayden-ledbetter M;

WPI; 2002-599691/64.

New human binding domain-immunoglobulin fusion protein useful for treating a subject having or suspected of having a B-cell disorder or malignant condition e.g. rheumatoid arthritis -

Claim 10; Page 79; 136pp; English.

The invention describes a binding domain-immunoglobulin fusion protein that is capable of at least one immunological activity, comprising a binding domain polypeptide fused to an immunoglobulin hinge region polypeptide capable of specifically binding to an antigen, or an immunoglobulin heavy chain CH2 or CH3 constant region polypeptide fused to the hinge region polypeptide or to the CH2 constant region polypeptide. The fusion protein is useful for treating a subject having or suspected of having a B-cell disorder or malignant condition e.g.

CC rheumatoid arthritis, myasthenia gravis, Grave's disease, Hashimoto's thyroiditis, type I diabetes mellitus, multiple sclerosis, systemic lupus erythematosus, Sjogrens syndrome, immune thrombocytopenic purpura, psoriasis, scleroderma, cancer and inflammatory bowel disease such as Chron's disease and ulcerative colitis. The fusion protein retains the ability to participate in well known immunological effector activities including antibody dependent cell mediated cytotoxicity and/or complement fixation in complement dependent cytotoxicity, despite having structures that would not be expected to be capable of promoting the effector activities. It can be produced in substantial quantities that are typically greater than those routinely attained with single-chain antibody constructs. This sequence represents a linker peptide used in the creation of binding domain-immunoglobulin fusion proteins.

Sequence 5 AA;

Query Match 100.0%; Score 5; DB 23; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGS 5
|||||

Db 1 GGGS 5
|||||

RESULT 41

ABG31505
ID ABG31505 standard; Peptide; 5 AA.

XX AC
XX AC
XX AC

05-NOV-2002 (first entry)

Flexible peptide linker #1.

Interferon gamma; IFNG; transforming growth factor beta; TGFb; latency; latency associated peptide; LAP; antiinflammatory; osteopathic; antiarthritic; dermatological; antirheumatic; cancer; neuroprotective; antiatherosclerotic; inflammatory disorder; osteoarthritis; scleroderma; renal disease; rheumatoid arthritis; inflammatory bowel disease; multiple sclerosis; atherosclerosis; complement fixation.

Synthetic.

WO200255098-A2.

18-JUL-2002.

09-JAN-2002; 2002WO-GB00068.

09-JAN-2001; 2001GB-0000551.

(QUEE-) QUEEN MARY & WESTFIELD COLLEGE.

Chernajovsky Y, Dreja HS, Adams G;

WPI; 2002-575421/61.

Providing latency to a pharmaceutically active agent, involves associating a fusion protein comprising a latency associated peptide and a proteolytic cleavage site with the pharmaceutically active agent -

Example 1; Page 27; 65pp; English.

This invention relates to a novel method for providing latency to a pharmaceutically active agent. The method involves associating a fusion protein comprising a latency associated peptide (LAP) and a proteolytic cleavage site with the pharmaceutically active agent. A pharmaceutical agent of the invention may have antiinflammatory, osteopathic, antiarthritic, dermatological, antirheumatic, neuroprotective and antiatherosclerotic activities. The method of the invention is useful for providing latency to a pharmaceutically active agent. A nucleic acid

CC encoding the fusion protein of the invention is useful for treatment of
 CC a patient such as a human or non-human animal by administering to the
 CC patient a therapeutically effective amount of the nucleic acid where the
 CC treatment is the treatment of an inflammatory disorder or gene therapy.
 CC The compositions of the invention are useful in medicine or in the
 CC manufacture of a medicament for the treatment of an inflammatory
 CC disorder. A fusion protein of the invention is useful for treating a
 CC patient, for use in medicine or in the manufacture of a medicament for
 CC the treatment of an inflammatory disorder. The fusion protein is also
 CC useful for providing latency to a pharmaceutically active agent or for
 CC providing site-specific activation of a latent biologically active
 CC agent. The method of the invention is useful for overcoming toxic
 CC effects of systemic administration of potent biological agents. The
 CC fusion protein or a nucleic acid encoding it is useful in veterinary
 CC medicine for treatment/prophylaxis of domestic animals including horses
 CC and companion animals (e.g. cats and dogs) and farm animals.
 CC Compositions of the invention are also useful for treating inflammatory
 CC disorder such as osteoarthritis, scleroderma, renal disease, rheumatoid
 CC arthritis, inflammatory bowel disease, multiple sclerosis,
 CC atherosclerosis, or cancer. The present sequence represents a flexible
 CC peptide linker used in teh cloning of the fusion protein constructs
 CC of the invention.

SQ Sequence 5 AA;
 Query Match 100.0%; Score 5; DB 23; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
 |||||
 DB 1 GGGGS 5

RESULT 42

ABG31871
 ID ABG31871 standard; peptide; 5 AA.

XX
 AC ABG31871;
 XX
 DT 05-NOV-2002 (first entry)
 XX
 DE Spacer peptide used to link VH and VL domain of single chain antibody.

XX Transgenic; disease-resistant trait.

XX Synthetic.

XX WO200229032-A2.

XX PD 11-APR-2002.

XX 01-OCT-2001; 2001WO-US31004.

XX 30-SEP-2000; 2000US-0677584.

XX 28-MAR-2001; 2001US-279702P.

XX 14-JUN-2001; 2001WO-US19367.

XX (DIVE-) DIVERSA CORP.

XX Short JM, Fu P, Latterich M, Wei J, Levin M;

XX WPI; 2002-618960/66.

XX Producing improved organism having a desirable trait, involves
 PT generating a set of mutagenised organisms so that a set of substantial
 PT genetic mutations is represented and detecting presence of improved
 PT organism -

XX Disclosure; Page 63; 869pp; English.

XX The invention relates to a method of producing an improved organism
 CC having a desirable trait, comprising obtaining an initial population of

CC organisms, generating a set of mutagenised organisms, so that when all
 CC the genetic mutations in the set of mutagenised organisms are taken as a
 CC whole, there is represented a set of substantial genetic mutations, and
 CC detecting the presence of the improved organism. Alternatively, the
 CC method involves functionally knocking out an endogenous gene in a clonal
 CC population of organisms, transferring a library of altered genes into the
 CC clonal population of organism, where each altered gene differs from the
 CC endogenous gene at only one codon, detecting a mutagenised organism
 CC having an improved trait, and determining the nucleotide sequence of a
 CC gene that has been transferred into the detected organism. The methods
 CC are useful for producing an improved organism having a desirable trait.
 CC The methods are also useful for defining the expressed proteins
 CC associated with a given cellular state, and quantifying changes in
 CC protein expression between at least two cellular states. The methods are
 CC also useful for simultaneously identifying individual proteins in
 CC complex mixtures of biological molecules. The methods are useful in
 CC functional genomic studies, and for conferring disease-resistant traits
 CC and traits other than disease resistance in plants. The present
 CC sequence represents a spacer peptide used to link VH and VL domains of a
 CC single chain antibody used in the method of the invention.

SQ Sequence 5 AA;

Query Match 100.0%; Score 5; DB 23; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
 |||||
 DB 1 GGGGS 5

RESULT 43

ABP52872
 ID ABP52872 standard; peptide; 5 AA.

XX AC ABP52872;

XX DT 03-NOV-2002 (first entry)

XX Flexible peptide linker SEQ ID NO:1.

XX Flexible peptide linker; alliinase; alliin; allacin; cytostatic;
 KW enzyme therapy; cancer; infectious disease.

XX Synthetic.

XX WO200258624-A2.

XX PD 01-AUG-2002.

XX 26-DEC-2001; 2001WO-US49384.

XX 26-DEC-2000; 2000IL-0140555.

XX (VEDA) YEDA RES & DEV CO LTD.

XX (MCIN/) MCINNIS P.

XX Rabinkov A, Miron T, Mirelman D, Wilchek M;

XX WPI; 2002-627377/67.

XX New alliinase conjugate in an enzymatically inactive form with a
 PT targeting carrier that guides the enzyme to a cell or to a
 PT microorganism of interest in the body, useful for treating cancer and
 PT infectious diseases -

XX Disclosure; Page 23; 70pp; English.

XX The present invention describes a conjugate of the enzyme alliinase in
 CC an enzymatically inactive form with a targeting carrier that guides the
 CC enzyme to a cell or to a microorganism of interest in the body. Also
 CC described: (1) a pharmaceutical composition comprising the conjugate

defined above, and a pharmaceutical carrier; (2) a pharmaceutical kit comprising in separate compartments a pharmaceutical composition of (1), and a composition of alliin; and (3) treating a disorder or disease treatable with alliin, comprising administering to an individual a conjugate of the enzyme alliinase with a targeting carrier that targets the conjugate to a desired tissue or organ in the body followed by administration of alliin, where alliin is generated at the tissue or organ to exert its biological activity. The conjugate has cytostatic activity and can be used in enzyme therapy. The alliinase conjugate is useful for treating several types of cancers and infectious diseases. The present sequence represents an example of a flexible peptide linker which is given in the exemplification of the present invention.

CC Sequence 5 AA;

Query Match 100.0%; Score 5; DB 23; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

AY 1 GGGS 5
|
|
|
|
|
b 1 GGGS 5

RESULT 44

AAO22207
ID AAO22207 standard; Peptide; 5 AA.

AC AAO22207;

DT 11-OCT-2002 (first entry)

Zinc finger protein peptide linker SEQ ID No 6.

Non-canonical zinc finger binding protein; ZFP; gene therapy.

Unidentified.

WO200257293-A2.

25-JUL-2002.

22-JAN-2002; 2002WO-US01893.

22-JAN-2001; 2001US-263445P.

11-MAY-2001; 2001US-290716P.

(SANG-) SANGAMO BIOSCIENCES INC.

Rebar E, Jamieson A;

WPI; 2002-566791/60.

Non-canonical zinc finger binding protein for modulating gene expression comprises non-canonical zinc finger components that bind to a target sequence -

Disclosure; Page 19; 63pp; English.

The invention relates to an isolated, non-canonical (e.g., non-C2H2) zinc finger binding protein (ZFP) comprising one or more non-canonical zinc finger components that bind to a target sequence. A fusion polypeptide of the invention is useful for modulating expression of a gene. The non-canonical ZFP and its encoding polynucleotide, and a fusion protein comprising the non-canonical ZFP and its encoding polynucleotide can be used to treat disease. The non-canonical ZFP can be used in diagnostic assays and to link phenotype to expression of particular genes. The polynucleotide encoding the non-canonical ZFP can be used to treat disorders by gene therapy. This sequence represents a peptide relating to the zinc finger binding protein of the invention.

CC Sequence 5 AA;

Query Match 100.0%; Score 5; DB 23; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGS 5
|
|
|
|
|
Db 1 GGGS 5

RESULT 45

ABG31102
ID ABG31102 standard; peptide; 5 AA.

AC ABG31102;

DT 21-OCT-2002 (first entry)

Linker peptide for glucocerebrosidase fusion proteins #7.

Glucocerebrosidase; GCR; fusion protein; IGG1; IGG2; linker peptide; Fc; glycolipid storage disorder; Gaucher's disease; Fabry's disease; Tay-Sachs disease; nephrotropic; antileptic.

Synthetic.

WO200257435-A2.

25-JUL-2002.

27-DEC-2001; 2001WO-EP15328.

18-JAN-2001; 2001EP-0101056.

(MERE) MERCK PATENT GMBH.

Schumacher S, Gillies S;

WPI; 2002-599719/64.

New fusion proteins with glucocerebrosidase activity, useful for treating glycolipid storage disorders, specifically Gaucher's disease, Fabry's disease or Tay-Sachs disease -

Disclosure; Page 9; 25pp; English.

The invention relates to a fusion protein, which comprises essentially of an immunoglobulin molecule (ig e.g. IgG1 Fc or IgG2 Fc) or its fragment and a non-immunoglobulin molecule. The non-immunoglobulin molecule is a protein having the biological activity of glucocerebrosidase (GCR-like protein). Also included are a DNA sequence encoding the fusion protein comprising: (a) a signal/leader sequence; (b) an Ig molecule; or (c) a target protein sequence having the biological activity of GCR. Also included are an expression vector comprising the fusion protein DNA, a host cell for expressing the novel fusion protein and comprising the vector, producing the novel fusion protein and a pharmaceutical composition comprising the novel fusion protein, and at least one pharmaceutical carrier, diluent or excipient. The fusion protein is useful for treating glycolipid storage disorders, specifically Gaucher's disease, Fabry's disease or Tay-Sachs disease. The fusion protein is also useful for the manufacture of a pharmaceutical composition for treating these diseases. The present sequence represents a linker peptide suitable for inclusion in a fusion protein of the invention.

CC Sequence 5 AA;

Query Match 100.0%; Score 5; DB 23; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGS 5
|
|
|
|
|
Db 1 GGGS 5

```

RESULT 46
AAO22078
ID AAO22078 standard; Peptide; 5 AA.
XX
XX AAO22078;
AC
XX 25-SEP-2002 (first entry)
DT
XX
XX 5 amino acid flexible linker of FKBP and FRB dimerisation system.
DE
XX
XX Protein fragment complement assay; PCA; penicillin; cephalosporin; TEM-1;
KW beta-lactamase; linker; BLF fragment; GCN 4; FKBP/FRB domain.
XX
XX Unidentified.
OS
XX WO200194617-A2.
XX
XX 13-DEC-2001.
PD
XX
XX 01-JUN-2001; 2001WO-US17886.
PF
XX
XX 02-JUN-2000; 2000US-208485P.
PR
XX 31-MAY-2001; 2001US-0870018.
XX
XX (ODYS-) ODYSSEY PHARM INC.
PA
XX
XX Michnick SW, Galarneau A;
PI
XX WPI; 2002-566445/60.
DR
XX
XX Protein fragment complementation assay, useful for the detection of
PT protein-protein interaction, protein-small molecule interaction and
PT protein-nucleic acid interaction, based on Escherichia coli TEM1
PT beta-lactamase -
XX
XX
PS Example 1; Page 9; 29pp; English.
XX
XX The invention relates to a protein fragment complement assay (PCA)
CC comprising allowing at least two molecules capable of mutual interaction
CC to draw into close molecular proximity at least two reporter molecule
CC fragments which, when in close molecular proximity, forms a complex
CC capable of reaction with a penicillin- or cephalosporin-class substrate,
CC and detecting a signal resulting from the reaction. The method is useful
CC for the detection of protein-protein, protein-small molecule and protein-
CC nucleic acid interactions, based on the Escherichia coli TEM-1 beta-
CC lactamase. This sequence represents a 5 amino acid flexible linker of
CC FKBP and FRB dimerisation system relating to the invention.
XX
XX
SQ Sequence 5 AA;
Query Match 100.0%; Score 5; DB 23; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
Db 1 GGGGS 5

RESULT 47
ABB83546
ID ABB83546 standard; peptide; 5 AA.
XX
XX ABB83546;
AC
XX 27-SEP-2002 (first entry)
DT
XX
XX Peptide linker #2.
DE
XX
XX Zinc finger; stress tolerance; pathogen resistance;
KW agrochemical; peptide linker.
XX

```

```

OS Unidentified.
XX
XX Key Location/Qualifiers
FH Misc-difference 1..5
FT /note= "Sequence can be optionally repeated n times"
FT
XX
XX WO200257294-A2.
XX
XX 25-JUL-2002.
PD
XX
XX 22-JAN-2002; 2002WO-US01906.
PF
XX
XX 22-JAN-2001; 2001US-263445P.
PR
XX 11-MAY-2001; 2001US-290716P.
XX
XX (SANG-) SANGAMO BIOSCIENCES INC.
PA
XX
XX Jamieson A, Li G;
PI
XX WPI; 2002-566792/60.
DR
XX
XX Modified plant zinc finger protein for modulating gene expression in a
PT plant cell comprises zinc fingers that bind to a target site -
XX
XX Disclosure; Page 22; 50pp; English.
XX
XX The present invention relates to a modified plant zinc finger
CC protein. This zinc finger protein is used to modulated gene
CC expression in a plant cell. Nucleic acid encoding the zinc finger is
CC expressed in plant cells to produce a plant with an altered phenotype
CC relative to the wild-type plant. The altered phenotype is high in
CC nutritional value, yield, stress tolerance, pathogen resistance,
CC resistance to agrochemicals, production of pharmaceutical compounds or
CC production of industrial chemicals. The present sequence is
CC a peptide linker used to link the zinc finger to zinc finger
CC proteins.
XX
XX
SQ Sequence 5 AA;
Query Match 100.0%; Score 5; DB 23; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
Db 1 GGGGS 5

RESULT 48
ABJ03945
ID ABJ03945 standard; Peptide; 5 AA.
XX
XX ABJ03945;
AC
XX 25-SEP-2002 (first entry)
DT
XX
XX Peptide linker SEQ ID NO: 212.
DE
XX
XX Zinc finger protein; angiogenesis; vasculogenesis; ischaemia;
KW diabetic retinopathy; psoriasis; arthropathy; cancer; tumour growth;
KW gene therapy; antiatherosclerotic; vasotropic; antithrombotic; vulnery;
KW antiulcer; cytostatic; antipsoriatic; antidiabetic; ophthalmological;
KW osteopathic; antiinfertility.
XX
XX Synthetic.
OS
XX
XX WO200246412-A2.
XX
XX 13-JUN-2002.
PD
XX
XX 06-DEC-2001; 2001WO-US46861.
PF
XX
XX 07-DEC-2000; 2000US-0733604.
PR

```

12-DEC-2000; 2000US-0736083.
30-APR-2001; 2001US-0846033.

(SANG-) SANGAMO BIOSCIENCES INC.

Rebar E, Jamieson A, Liu Q, Liu P, Wolffe A, Eisenberg SP,
Jarvis E;
WPI; 2002-527918/56.

New zinc finger protein that binds to target site in vascular
endothelial growth factor gene, useful for modulating expression of the
gene and for treating atherosclerosis, ischemia, arthritis, wound or
ulcer -

Disclosure; Page 187; 195pp; English.

The present invention relates to a zinc finger protein that binds to a
target site in one or more vascular endothelial growth factor (VEGF)
genes. The protein is useful for modulating expression of a VEGF gene,
thereby regulating angiogenesis and vasculogenesis. This can be used to
treat atherosclerosis, ischemia, arthritis, wounds, ulcers, tumours,
diabetic retinopathy or psoriasis. The present sequence is a peptide
shown in the invention.

Sequence 5 AA;

Query Match 100.0%; Score 5; DB 23; Length 5;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GGGGS 5

|||||

1 GGGGS 5

RESULT 49

BP48193

ABP48193 standard; Peptide; 5 AA.

ABP48193;

28-AUG-2002 (first entry)

Peptide linker SEQ ID NO:3.

Zinc finger protein; ZFP; DNA binding protein; zinc finger.

Synthetic.

WO200242459-A2.

30-MAY-2002.

20-NOV-2001; 2001WO-US43438.

20-NOV-2000; 2000US-0716637.

(SANG-) SANGAMO BIOSCIENCES INC.

Liu Q;

WPI; 2002-500284/53.

New zinc finger protein that binds to target site, useful in studying
gene function and for human therapeutics and plant engineering,
comprises first, second and third zinc fingers, ordered from N- to
C-terminus -

Disclosure; Page 15; 81pp; English.

The present invention describes a zinc finger protein (I) that binds to
a target site, comprising a first (F1), a second (F2), and a third (F3)

zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
target site comprises, in 3'-5' direction, a first (S1), a second (S2),
and a third (S3) target subsite. Also described are: (I), a polypeptide
(II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
(3) designing (M) (I) involves selecting the F1 zinc finger such that
it binds to the S1 target subsite, selecting the F2 zinc finger such
that it binds to the S2 target subsite, and selecting the F3 zinc
finger such that it binds to the S3 target subsite, thus designing (I)
that binds to a target site. (I) is useful for recognition of triplet
target subsites having the nucleotide G in the 5'-most position of the
subsite. (I) is useful in studying gene function, and for human
therapeutics and plant engineering. (II), (II) or (III) is useful in
therapeutic methods to modulate the expression of a target region within
a subject, in diagnostic methods for sequence specific detection of
target nucleic acid in a sample, and in assays to determine the
phenotype and function of gene expression. (I) has improved affinity
and specificity for their target sequences, as well as enhanced
biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
represent DNA target sequences and zinc finger peptides which are given
in the exemplification of the present invention.

Sequence 5 AA;

Query Match 100.0%; Score 5; DB 23; Length 5;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GGGGS 5

|||||

1 GGGGS 5

RESULT 50

ABG60659

ID ABG60659 standard; Peptide; 5 AA.

AC ABG60659;

13-AUG-2002 (first entry)

Polyimmunoglobulin receptor (pIGR) associated linker sub-unit.

Transcellular transport; transcytotic transport; paracellular transport;
respiratory system disorder; lung cancer; tumour; asthma;
pathogenic infection; allergy-related disorder;

Gastrointestinal tract disorder; gastrointestinal hormone disorder;

Chron's disease; eating disorder; polyimmunoglobulin receptor; pIGR.

Synthetic.

WO200228408-A2.

11-APR-2002.

02-OCT-2001; 2001WO-US30832.

02-OCT-2000; 2000US-237929P.

13-NOV-2000; 2000US-248478P.

14-NOV-2000; 2000US-248819P.

09-FEB-2001; 2001US-267601P.

(ARIZ-) ARIZEKE PHARM INC.

Houston LL, Sheridan PJ, Hawley S, Glynn JM, Chapin S, Basu A;

WPI; 2002-416628/44.

Complex useful for transporting active agent through epithelial
barrier, has biologically active portion and target element directed to
ligand that confers e.g. transcytotic properties to agent specific to
ligand -

Example 6; Page 222; 379pp; English.

XX The invention described a complex or compound (I) comprising a
 CC biologically active portion and a target element (II) directed to a
 CC ligand that confers transcellular, transcytotic or paracellular
 CC transporting properties to an agent specifically bound to the ligand,
 CC where (II) is not an antibody. Alternatively, (I) comprises two or more
 CC (II) directed to one or more ligands. (I) is useful for delivering a
 CC biologically active agent to an animal, for transporting an active agent
 CC through an epithelial or mucosal barrier, and for treating or
 CC identifying a disease in an animal e.g. diseases of the respiratory
 CC system including lung cancer and tumours, asthma, pathogenic infections,
 CC allergy-related disorders, gastrointestinal tract disorders, disorders
 CC relating to gastrointestinal hormones, Chron's disease, eating disorders
 CC and any disease or disorder involving polyimmunoglobulin receptor (pigr)
 CC displaying cells. This sequence represents a peptide associated with the
 CC transport of biologically active agents across cellular barriers.
 XX
 SQ Sequence 5 AA;

Query Match 100.0%; Score 5; DB 23; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
 Db 1 GGGGS 5

RESULT 51

AAE22204
 ID AAE22204 standard; peptide; 5 AA.

AC AAE22204;

DT 25-JUL-2002 (first entry)

DE Peptide linker #2 used in construction of bispecific antibody CCR5xCD3.
 XX Chemokine construct; human immunodeficiency virus 1; allergic disease;
 KW skin disease; immunological disorder; autoimmune disease; psoriasis;
 KW multiple sclerosis; rheumatoid arthritis; inflammatory bowel disease;
 KW diabetes; skin inflammation; atopic dermatitis; inflammatory disease;
 KW inflammatory renal disease; HIV-1; transplant rejection; CCR5xCD3;
 KW antibody; linker.

XX Unidentified.

XX WO200220615-A2.

PN 14-MAR-2002.

PD 10-SEP-2001; 2001WO-EP10433.

PF 08-SEP-2000; 2000EP-0119694.

PR 05-SEP-2001; 2001US-0948004.

XX (MICR-) MICROMET AG.

XX Mack M, Schloendorff D, Spring M;

XX WPI; 2002-362240/39.

XX Use of an antibody and/or chemokine construct that binds to a chemokine
 CC receptor, for eliminating cells latently infected with primate
 CC immunodeficiency virus, or treating, preventing and alleviating immune
 CC disorders
 PT
 XX Disclosure; Page 51; 117pp; English.

XX The invention relates to the use of an antibody and/or chemokine
 CC construct that binds a chemokine receptor for preparing a pharmaceutical
 CC composition for eliminating cells latently infected with a primate
 CC immunodeficiency virus (e.g. human immunodeficiency virus (HIV-1)). They

CC are used in gene therapy and as vaccines. The antibody and/or chemokine
 CC construct is also used for preparing a pharmaceutical composition for
 CC treating, preventing and/or alleviating immunological disorders
 CC including autoimmune diseases (e.g. multiple sclerosis, type 1 diabetes
 CC and rheumatoid arthritis), allergic diseases, skin diseases (e.g. skin
 CC inflammation, atopic dermatitis and psoriasis), inflammatory diseases
 CC such as inflammatory joint diseases (chronic arthritis), inflammatory
 CC renal diseases and inflammatory bowel diseases and graft versus host
 CC and transplant rejections. The present sequence is a linker peptide
 CC used in the construction and expression of bispecific single chain
 CC antibody CCR5xCD3 construct of the invention.

XX Sequence 5 AA;

Query Match 100.0%; Score 5; DB 23; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
 Db 1 GGGGS 5

RESULT 52

ABB75959
 ID ABB75959 standard; Peptide; 5 AA.

XX ABB75959;

DT 12-JUL-2002 (first entry)

XX Linker peptide.

XX Linker: 4-LBB ligand.

XX Synthetic.

XX US6355779-B1.

PN 12-MAR-2002.

PD 10-SEP-1998; 98US-0150864.

XX 06-MAY-1994; 94US-0236918.

PR 05-AUG-1997; 97US-0910449.

PR 07-MAY-1993; 93US-0060843.

XX (IMMV) IMMUNEX CORP.

XX Goodwin RG, Smith CA, Alderson MR;

XX WPI; 2002-380940/41.

XX New antibody specific for the cytokine 4-LBB-ligand, useful for
 CC immunoadfinity purification of the ligand
 PT
 XX Disclosure; Column 7; 31pp; English.

XX The present sequence is a linker peptide that can be used in fusion
 CC proteins of the invention. The linker sequence is preferably
 CC repeated 4-12 times. The invention provides novel murine and
 CC human 4-LBB ligand (4-LBB-L) cytokine polypeptides and human 4-LBB
 CC polypeptides, as well as DNA sequences encoding them, recombinant
 CC expression vectors and host cells, and methods for producing the
 CC novel polypeptides by cultivating the transformed host cells.
 CC Soluble forms of the 4-LBB-L and 4-LBB polypeptides, derived from
 CC their extracellular domains, have therapeutic value.

XX Sequence 5 AA;

Query Match 100.0%; Score 5; DB 23; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGGGS 5
b |||||
1 GGGGS 5

RESULT 53
AAE21141
D AAE21141 standard; peptide; 5 AA.

X X
X C
X C
X X
X T
X T
X T
E Linker peptide, to join multi-finger ZFPs to other multi-finger ZFPs.
X DNA binding protein; zinc finger domain; zinc finger protein; ZFP;
X viral replication; gene expression; virucide.
X Synthetic.
X WO200208286-A2.
X 31-JAN-2002.
X 19-JUL-2001; 2001WO-EP08367.
X 21-JUL-2000; 2000US-220060P.
X (SYGN) SYNGENTA PARTICIPATIONS AG.
X Sera T;
X WPI; 2002-172000/22.

New Zinc Finger Protein (ZFP) comprising three essential domains useful
for diagnosing diseases associated with abnormal genomic structure -
Disclosure; Page 59; 143pp; English.

The present invention relates to novel DNA binding proteins comprising
zinc finger domains in which two histidine and two cysteine residues
coordinate a central zinc ion. The invention particularly relates to
the identification of a context-independent recognition code to design
zinc finger domains. The invention also relates to zinc finger proteins
(ZFP) designed using this recognition code. The ZFPs are useful for
altering genomic structure, inhibiting viral replication (where viral
replication is inhibited for plant virus, an animal virus or a human
virus), modulating gene expression, detecting an altered zinc finger
recognition sequence and diagnosing disease associated with abnormal
genomic structure. The present sequence is a linker peptide which is
used to join multi-finger ZFPs with other multi-finger ZFPs.

XX Sequence 5 AA;
Query Match 100.0%; Score 5; DB 23; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Yy 1 GGGGS 5
b |||||
1 GGGGS 5

RESULT 54
AAO17155
D AAO17155 standard; Peptide; 5 AA.

XX X
X C
X C
X T
X T
X T
DE Human endoglin (CD105) binding protein related linker peptide A.

XX Human endoglin; CD105; binding protein; cancer; cytostatic;
KW neovascularisation; linker.
XX Synthetic.
XX WO200220614-A2.
XX 14-MAR-2002.
XX 04-SEP-2001; 2001WO-EP10197.
XX 04-SEP-2000; 2000DE-1043481.
XX (VECT-) VECTRON THERAPEUTICS AG.
XX Kontermann R, Miller D, Mueller R;
PI WPI; 2002-292258/33.
XX New polypeptide that binds endoglin, useful for diagnosis and treatment
PT of tumors, is synthetic single-chain antibody -
XX Example 2; Page 19; 37pp; German.
XX The present invention relates to a polypeptide capable of binding
CC specifically to the extracellular domain of human endoglin (CD105). This
CC can be used for the detection of CD105 or cells that express CD105 and
CC to bind to CD105-expressing cancer cells to produce a cytotoxic effect,
CC i.e. functioning as a carrier for active agents such as adenoviruses or
CC cytotoxic T lymphocytes. In particular, it can be used in the diagnosis
CC and therapy of diseases associated with hyperproliferation of CD105+
CC cells, especially tumours. The present sequence is a linker peptide
CC used in the exemplification of the invention.

XX SQ Sequence 5 AA;
Query Match 100.0%; Score 5; DB 23; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db |||||
1 GGGGS 5

RESULT 55
AAE19822
ID AAE19822 standard; peptide; 5 AA.

XX AC AAE19822;
XX DT 18-JUN-2002 (first entry)

XX Linker peptide #1 used in the exemplification of the invention.
KW Expression system; interleukin-12; cancer therapy; cellular vaccine;
KW immunotherapeutic response; IL-12.4; linker.

XX OS Synthetic.
XX US2002018767-A1.
XX 14-FEB-2002.

XX 10-APR-2001; 2001US-0828825.
XX 27-JUL-2000; 2000KR-0043498.

XX (LEES/) LEE S.
XX (KIMH/) KIM H.
XX Lee S, Kim H;
PI

XX
DR WPI; 2002-239247/29.
XX
XX New expression system, useful as vaccine for treating cancer, comprises
PT DNA molecule encoding interleukin-12 and costimulatory molecule,
PT operably linked to promoter that directs the expression of DNA molecule
PT
XX
XX Example 1; Page 5; 76pp; English.
XX
XX The invention relates to an expression system which comprises a DNA
CC molecule encoding interleukin (IL)-12 and a costimulatory molecule,
CC operably linked to a promoter that directs the expression of the DNA
CC molecule. Expression vectors of the invention are useful for in vitro
CC generation of genetically modified human cancer cells for cancer
CC therapy. These cells share phenotypes of both antigen presenting cells
CC and cancer cells and are suitable as a cellular vaccine for certain
CC types of cancer. Expression vectors are useful for eliciting an anti-
CC tumour immune response in a individual. They are useful for activating
CC T cells for immunotherapeutic responses against primary or metastatic
CC cancers. The present sequence is a linker peptide used in the
XX exemplification of the invention.
XX
SQ Sequence 5 AA;
Query Match 100.0%; Score 5; DB 23; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGGS 5
DB 1 GGGGS 5
RESULT 56
AAE19513
ID AAE19513 standard; peptide; 5 AA.
XX
AC AAE19513;
XX
DT 31-MAY-2002 (first entry)
XX
DE Linker peptide used to generate scFv antibody.
XX
KW Recombinant DNA construct; secretory leader sequence; His tag; infection;
KW therapy; RGDp; secreted replicable genetic display package; vaccination;
KW restriction endonuclease; cleavage site; phage display library;
KW antibacterial.
XX
OS Unidentified.
XX
PN WO200212513-A2.
XX
XX 14-FEB-2002.
XX
XX 24-JUL-2001; 2001WO-GB03328.
XX
XX 04-AUG-2000; 2000GB-0019049.
XX
XX (NEUT-) NEUTECH PHARMA PLC.
XX
XX Burnie JP, Matthews RC, Rigg G;
XX
XX WPI; 2002-227158/28.
XX
XX Novel recombinant DNA construct comprising DNA sequence for secreted
PT replicable genetic display package out of frame with DNA sequences for
PT secretory leader sequence, tag, and two restriction enzyme recognition
PT sites -
XX
XX Disclosure; Page 19; 54pp; English.
XX
XX The present invention relates to a recombinant DNA construct coding for

CC secretory leader sequence, tag, at least two restriction endonuclease
CC cleavage sites, and a component of a secreted replicable genetic display
CC package (RGDP) which, when expressed in frame in a recombinant host
CC organism, causes the RGDP to display at the surface of the package, the
CC tag and any polypeptide encoded 3' of the tag. The method of the
CC invention is also useful for producing a member of SBP (specific binding
CC pair) where the recombinant DNA construct is expressed in recombinant
CC host organism or recombinant host cells. The recombinant host organism or
CC host cells are infected with a helper phage and the method involves
CC isolating the Kbfra displaying the tag, the method is useful for
CC manufacturing a phage display library where several recombinant DNA
CC constructs prepared using several DNA sequences for components of
CC specific binding pairs, are employed. Members (antibodies or their
CC antigen binding fragments, or scFv fragments) of SBP identified have
CC therapeutic or diagnostic applications. The coding sequence may be
CC determined and synthesised in bulk for use as diagnostic or therapeutic
CC agent for treating infection. The SBP members are also useful for
CC isolating and purifying their complementary binding partners which are
CC used in vaccination. They are also used in epitope mapping, etc. The SBP
CC sequences are used for preparing a library of database of sequences, e.g.
CC the CDR3 region of antibodies and antigen binding fragments are useful
CC for determining the identity of the antigen bound by the antibody or its
CC fragment. This library can be examined to determine the identity of
CC conserved sequences and to identify members of SBPs which are good
CC therapeutic or diagnostic agents. The present sequence is a linker
CC peptide used in the generation of human scFv antibodies.
XX
SQ Sequence 5 AA;
Query Match 100.0%; Score 5; DB 23; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGGS 5
DB 1 GGGGS 5
RESULT 57
AAO14387
ID AAO14387 standard; Peptide; 5 AA.
XX
XX AAO14387;
XX
XX 24-MAY-2002 (first entry)
XX
XX Linker peptide for use in a fusion protein #1.
XX
XX Human; metalloproteinase-like/disintegrin-like protein; MPD;
KW angiogenesis; endothelial cell migration; integrin ligand binding;
KW ocular disorder; cancer; inflammation; osteoporosis; restenosis;
KW cytostatic; antiinflammatory; osteopathic; vasotropic; thrombolytic;
KW vulnary; antiallergic; antibacterial; virucide; protozoacide; cardiant;
KW antianaemic; analgesic; thrombosis; tissue repair; neural disorder;
KW infection; linker.
XX
XX Synthetic.
OS
XX WO200210406-A2.
XX
XX 07-FEB-2002.
XX
XX 27-JUL-2001; 2001WO-US23734.
XX
XX 28-JUL-2000; 2000US-221838P.
XX
XX (IMV) IMMUNEX CORP.
XX
XX DuBoose RF, Wiley SR, Black RA;
XX WPI; 2002-195965/25.
XX
XX Novel purified human metalloproteinase-disintegrin polypeptide, useful

T for treating ocular and inflammatory disorders, osteoporosis, cancer,
 T restenosis, thrombosis, and chronic pain conditions -
 X Disclosure; Page 20; 101pp; English.

X The present invention provides a number of human metalloproteinase-like
 C /disintegrin-like (MPD) proteins and their fragments. These can be used
 C to identify treatments for ocular disorders, malignant and metastatic
 C conditions, inflammatory diseases, osteoporosis and other conditions
 C mediated by accelerated bone resorption, restenosis, inappropriate
 C platelet activation, recruitment or aggregation, thrombosis, conditions
 C requiring tissue repair or wound healing, endothelial migration,
 C angiogenesis, allergies, reproductive, neurological and vascular
 C conditions, infections, cardiovascular disorders such as myocardial
 C infarction, chronic pain conditions, endocrine system disorders,
 C gastrointestinal system disorders, genitourinary system disorders,
 C anaemia and haematological disorders. The present sequence is a linker
 C peptide described in the exemplification of the invention.

X Q Sequence 5 AA;

Query Match 100.0%; Score 5; DB 23; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGGGS 5
 b 1 GGGGS 5

RESULT 58

AAU75495
 D AAU75495 standard; Peptide; 5 AA.

X C AAU75495;

T 08-MAY-2002 (first entry)

E Human Gly4Ser peptide linker.

X Human growth hormone; GH; GHR; Gly4Ser linker;
 W growth hormone; receptor; GHatopGHR SD100; GHlinkGHR; GHlinkGHRflec;
 W pRcRNSsaci; TrcHndrev; Chi 1A2 chimera; acromegaly; gigantism;
 W growth hormone deficiency; Turner's syndrome; renal failure;
 W osteoporosis; diabetes mellitus; cancer; obesity; insulin resistance;
 W hyperlipidaemia; hypertension; anaemia; autoimmune disease;
 W infectious disease; inflammatory disorder; rheumatoid arthritis;
 W interleukin-6 chimera.

X S Synthetic.

X WO200196565-A2.

X 20-DEC-2001.

X 18-JUN-2001; 2001WO-GB02645.

X 16-JUN-2000; 2000GB-0014765.

X 10-MAR-2001; 2001GB-0005969.

X 16-MAR-2001; 2001GB-0006487.

X (ASTE-) ASTERION LTD.

X Ross R, Artymiuik P, Sayers J;

X WPI; 2002-130734/17.

X New binding agent useful in producing a medicament for treating e.g.
 T cancer, obesity, acromegaly or gigantism, comprises a first part that
 T binds to a ligand binding domain of a receptor and a second part having
 T a receptor binding domain -

X Claim 21; Page 47; 79pp; English.

XX The invention relates to a binding agent comprising a first part capable
 CC of binding a ligand binding domain of a receptor linked to a second part
 CC comprising a receptor binding domain, where the binding agent modulates
 CC the activity of the receptor. Also included are a nucleic acid molecule
 CC having a sequence, which encodes a binding agent comprising sequences
 CC given in the specification comprising the sequences of the full length
 CC GHatopGHR SD100 construct, GHlinkGHR construct(GH, growth hormone,
 CC GHR, growth hormone receptor), GHlinkGHRflec construct, 1157 base pair
 CC PCR fragment GHlinkGHR generated by nucleotides pRcRNSsaci and
 CC TrcHndrev, or the nucleotide sequence of the Chi 1A2 chimera,
 CC sequences binding to the nucleic acids or degenerate sequences
 CC representing them (which have receptor antagonising activity),
 CC their encoded polypeptides, a vector comprising the nucleic acids and a
 CC cell transformed/transfected with the nucleic acid or vector.
 CC The binding agent is useful for manufacturing a medicament for
 CC the treatment of acromegaly, gigantism, growth hormone (GH) deficiency,
 CC Turner's syndrome, renal failure, osteoporosis, diabetes mellitus,
 CC cancer, obesity, insulin resistance, hyperlipidaemia, hypertension,
 CC anaemia, autoimmune and infectious diseases, and inflammatory disorders
 CC including rheumatoid arthritis (interleukin (IL)-6 chimera).
 CC The present sequence is the Gly4Ser peptide linker used to join the
 CC 2 parts of the fusion proteins of the invention

XX Sequence 5 AA;

Query Match 100.0%; Score 5; DB 23; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
 Db 1 GGGGS 5

RESULT 59

AAU76043

ID AAU76043 standard; Peptide; 5 AA.

AC AAU76043;

XX 08-MAY-2002 (first entry)

XX Synthetic linker moiety peptide sequence #1.

XX Human; A Disintegrin And Metalloprotease with Homology to ADAM9;
 KW ADAM-H9; cytostatic; antiinflammatory; osteopathic; vasotropic;
 KW thrombolytic; vulnary; antiallergic; antibacterial; virucide;
 KW prozoacide; analgesic; angiogenesis; endothelial cell migration;
 KW ocular disorder; restenosis; thrombosis; cardiovascular disorder;
 KW chronic pain condition; haematological disorder; cancer; tumour.

XX Synthetic.

OS WO200210405-A2.

PN 07-FEB-2002.

XX 27-JUL-2001; 2001WO-US23709.

XX 28-JUL-2000; 2000US-221838P.

PR 09-APR-2001; 2001US-282550P.

XX (IMMV) IMMUNEX CORP.

XX Black RA, Poindexter K, Mosley BA, Dubose RF, Wiley SR;

XX WPI; 2002-189750/24.

XX Novel substantially purified human polypeptide having homology to A
 PT Disintegrin and Metalloprotease domain (ADAM) polypeptide family,
 PT termed ADAM-H9, for treating ocular and inflammatory diseases,
 PT osteoporosis, cancer -

XX Disclosure; Page 13; 72pp; English.

XX The present invention relates to a new human polypeptide having homology

XX to A Disintegrin And Metalloproteinase domain (ADAM) polypeptide family,

CC termed ADAM-H9. The polypeptide of the invention is useful for

CC identifying an agent that modulates an activity of the polypeptide, for

CC inhibiting angiogenesis in a mammal, for modulating angiogenesis in a

CC tissue, for modulating endothelial cell migration in vitro or in vivo,

CC for inhibiting the binding of an integrin to a ligand and for modulating

CC the binding of an integrin to a ligand in a mammal afflicted with a

CC condition such as ocular disorders, malignant and metastatic conditions,

CC inflammatory diseases, osteoporosis, accelerated bone resorption

CC disorders, retinosis, inappropriate platelet activation, recruitment or

CC aggregation, thrombosis and a condition requiring tissue repair or wound

CC healing. The polypeptide used for treating the above conditions is a

CC multimer (a dimer or trimer) which comprises an FC polypeptide, a leucine

CC zipper or a peptide linker. The invention is useful for treating

CC disintegrin-associated disorders and conditions and is also useful for

CC treating medical conditions and diseases associated with cell-cell and

CC cell matrix interactions, endothelial migration, angiogenesis,

CC inflammation, allergy, reproductive, neurological and vascular

CC conditions, bacterial, viral or protozoal infections, cardiovascular

CC disorders such as myocardial infarction, chronic pain conditions,

CC endocrine system disorders, gastrointestinal system disorders,

CC genitourinary system disorders, anaemia, haematological disorders and

CC oncological conditions such as cancer. The present amino acid sequence

CC represents one of several (AAU76043-AAU76049) linker moiety peptides of

CC the invention. These linkers were used in the invention to join an

CC ADAM-H9 disintegrin domain to an FC domain or leucine zipper domain.

XX

SQ Sequence 5 AA;

Query Match 100.0%; Score 5; DB 23; Length 5;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5

DB 1 GGGGS 5

RESULT 60

AAU10629

ID AAU10629 standard; Peptide; 5 AA.

AC AAU10629;

XX

XX 25-JAN-2002 (first entry)

DE Antibody variable region fusion proteins, linker peptide.

XX

XX Antibody; heavy chain variable region; VH; light chain variable region;

KW VL; bioactivity; biomolecule; biological reagent; diagnostic;

KW therapeutic; polypeptide; flow cytometry; gene library.

XX

OS Synthetic.

XX

XX US2001041333-A1.

PN

XX

XX 15-NOV-2001.

PD

XX

XX 15-DEC-2000; 2000US-0738871.

PF

XX

XX 16-JUN-1997; 97US-0876276.

PR

XX

XX 16-JUN-1998; 98US-0098206.

PR

XX

XX 22-NOV-1999; 99US-044112.

PR

XX

XX 10-OCT-2000; 2000US-0685432.

PR

XX

XX (SHOR/) SHORT J M.

PA (KELL/) KELLER M.

XX

XX Short JM, Keller M;

PI

XX WPI; 2002-017215/02.

DR

XX

PT Identifying bioactivities or biomolecules in mixed populations of

PT organisms, useful e.g. for isolating polypeptide synthesis enzymes,

PT comprises screening mixed library with labelled probe -

XX

PS Disclosure; Page 28; 45pp; English.

XX

XX The invention describes a novel method of identifying a bioactivity or

CC biomolecule comprising: contacting a library of polynucleotide clones

CC derived from a mixed population of organisms with at least one

CC oligonucleotide probe labelled with a detectable molecule; and separating

CC clones with an analyser that detects the detectable molecule. The method

CC is particularly used to identify new enzymes, e.g. for development of

CC biological or diagnostic reagents, therapeutics or industrial compounds,

CC particularly enzymes involved in biosynthesis of polypeptides. It allows

CC high throughput screening of DNA, isolated directly from the environment

CC and isolating at least one organism from a sample, by flow cytometry,

CC using techniques developed for eukaryotic cells. Very rapid

CC identification of selected biological activities is achieved; libraries

CC can be produced from mixtures of hundreds of different microorganisms and

CC 30-200 million clones may be screened per hour. This sequence is an a

CC peptide linker used to join the C-terminus of a light chain variable

CC region (VL) to the N-terminus of a heavy chain variable region, or

CC vice-versa, to form antibody variable region fusion proteins described in

CC the method of the invention.

XX

SQ Sequence 5 AA;

Query Match 100.0%; Score 5; DB 23; Length 5;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5

DB 1 GGGGS 5

RESULT 61

AAU74892

ID AAU74892 standard; Peptide; 5 AA.

XX

AC AAU74892;

XX

XX 09-APR-2002 (first entry)

DT

XX

XX Complement pathway protein C3d, peptide mimetic spacer segment.

DE

XX

XX Complement; receptor; CD21; CD2; C3d; immune response;

KW B cell stimulator; vaccine; CD21/CD19 complex; tumour;

KW cancer.

XX

OS Synthetic.

XX

XX WO200192295-A2.

PN

XX

XX 06-DEC-2001.

PD

XX

XX 30-MAY-2001; 2001WO-CA00785.

PF

XX

XX 30-MAY-2000; 2000US-207434P.

PR

XX

XX (UTOR) UNIV TORONTO.

PA

XX

XX Iseman DE, Ciemenza L;

PI

XX

XX WPI; 2002-114323/15.

DR

XX

XX Ligand useful for modulating immune response such as in the preparation

PT of vaccine comprises CD21 contacting amino acid residues of C3d

PT molecule -

XX

S Disclosure; Page 19; 53pp; English.

X The invention describes a ligand of the complement receptor 2 (CD21 or CD2) comprising amino acid residues 36-39 and 160-167 of the C3d molecule. The ligand is useful in the manufacture of a medicament such as a vaccine for modulating the immune response of a host (preferably tumour vaccine), and as antigens in immunogenic compositions, therapeutics C diagnostic reagents, in the generation of diagnostic agents and as cancer C therapeutics. The ligand has the ability to bind CD21 and stimulate B C cells through the CD21/CD19 complex. Non-naturally occurring ligands C and site specific mutated analogues of C3d demonstrating an enhanced C binding affinity for CD21 as compared to the binding affinity of a C wild-type C3d molecule. The ligand alters the immunogenicity of an C antigen, e.g. by inducing or enhancing an immune response to an antigen C in a host and thus protects the host against disease caused by the C pathogen. This sequence represents a flexible spacer segment used to C link linear or cyclic peptide mimetics of C3d, a complement pathway C protein, described in the method of the invention.

X Sequence 5 AA;

Q Query Match 100.0%; Score 5; DB 23; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGGGS 5
b 1 GGGGS 5
|||||

ESULT 62
AAU75148
D AAU75148 standard; peptide; 5 AA.

X AAU75148;
C 23-APR-2002 (first entry)
T Single unit of linker peptide used in human antibody invention.
E Human antibody; immunoglobulin; Ig; variable heavy-chain; VH;
X variable light-chain; VL.
W Synthetic.
W WO200200729-A2.
N 03-JAN-2002.
D 25-JUN-2001; 2001WO-US20542.
F 23-JUN-2000; 2000US-0602373.
R 23-JUN-2000; 2000US-0602972.
R 23-JUN-2000; 2000US-0603658.
R 23-JUN-2000; 2000US-0603663.
A (GENE-) GENETASTIX CORP.
X Zhu L, Hua SB;
I WPI; 2002-090521/12.
X Screening libraries of tester proteins against protein, peptide or
R nucleic acid target(s) using a two-hybrid method in yeast, useful for
T generating recombinant human antibodies and screening for their
T affinity binding with target antigens -
X Claim 46; Page 170; 251pp; English.

S The present invention relates to compositions and methods for high
C throughput generation and screening of a human antibody or immunoglobulin
C (Ig) library in yeast. The method comprises expressing a library of
C tester fusion proteins in yeast cells, each tester fusion protein

CC comprising either an activation domain or a DNA binding domain of a
CC transcription activator and a tester protein having a large diversity
CC within the library. The tester protein comprises a first polypeptide
CC subunit (e.g. human variable heavy-chain, VH) whose sequence varies
CC within the library, a second polypeptide subunit (e.g. human variable
CC light-chain, VL) whose sequence varies within the library independently
CC of the first polypeptide, and a linker peptide which links the first and
CC second polypeptide subunits. The method is useful for generating
CC recombinant human antibodies and screening for their affinity binding
CC with target antigens. The present sequence represents the single unit of
CC the linker peptide used in the methods of the present invention.

XQ Sequence 5 AA;

Query Match 100.0%; Score 5; DB 23; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGGS 5
Db 1 GGGGS 5
|||||

RESULT 63
AAE34831
ID AAE34831 standard; peptide; 5 AA.

XX AAE34831;
AC AAE34831;
XX 28-MAY-2003 (first entry)
DT Peptide linker #1.
DE
XX
KW Siglec-12; therapy; organ transplant disorder; diabetes mellitus;
KW haematologic disorder; Panconi's aplastic anaemia; septic shock; stroke;
KW Alzheimer's disease; infection; asthma; cachexia; myocardial infarction;
KW immunosuppressive; thymometic; antibacterial; antiparasitic; cardiac;
KW liver disorder; cerebroprotective; infection; virucide; neuroprotective;
KW lymphoma; osteoporosis; dermatological; neotropic; rheumatoid arthritis;
KW dementia; protozoacide; vasotropic; Hashimoto's thyroiditis; cytostatic;
KW ischaemia; reperfusion injury; systemic lupus erythematosus; psoriasis;
KW osteopathic.
XX Unidentified.
OS WO200296452-A1.
PN 05-DEC-2002.
PD 29-MAY-2002; 2002WO-US16906.
XX 29-MAY-2001; 2001US-294199P.
XX (IMMV) IMMUNEX CORP.
PA Anderson DM, Marken JS;
PI WPI; 2003-140419/13.
DR Novel substantially purified Siglec-12 polypeptide useful for treating
XX rheumatoid arthritis, systemic lupus erythematosus, Hashimoto's
PT thyroiditis, asthma, Alzheimer's disease, dementia, cachexia,
PT myocardial infarction -
XX Disclosure; Column 16; 47pp; English.

PS The invention relates to Siglec-12 polypeptide and polynucleotide
XX sequences. Siglec-12 is useful for identifying an agent that modulates
CC its activity. The invention is useful for treating a Siglec-associated
CC disorder or disease (e.g. rheumatologic disorder, bone marrow or solid
CC organ transplant disorder, a graft-versus-host disorder, inflammatory
CC disorder, autoimmune disorder, neurologic disorder, cell proliferative
CC disorder, infection, cardiovascular disorder, haematologic disorder,

CC liver disorder or a bone disorder). Antibody that specifically binds
 CC to Siglec-12 is useful for treating a tumour that expresses Siglec-12
 CC polypeptide. The invention is also useful for treating systemic lupus
 CC erythematosus, rheumatoid arthritis, dementia, lymphoma, Hashimoto's
 CC thyroiditis, Alzheimer's disease, bacterial, parasitic, protozoal and
 CC viral infections, asthma, cachexia, myocardial infarction, ischaemia/
 CC reperfusion injury, stroke, diabetes mellitus, psoriasis, Fanconi's
 CC aplastic anaemia, septic shock, osteoporosis, etc. The present sequence
 CC is a peptide linker used in the invention.

XX SQ Sequence 5 AA;

Query Match 100.0%; Score 5; DB 24; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
 Db 1 GGGGS 5

RESULT 64

AAE35002
 ID AAE35002 standard; peptide; 5 AA.

XX AC AAE35002;

XX DT 28-MAY-2003 (first entry)

XX DE Linker peptide used to illustrate the method of the invention.

XX KW Variable domain; transcription-based amplification; cloning.

XX OS Unidentified.

XX PN WO200292770-A2.

XX PD 21-NOV-2002.

XX PF 14-MAY-2002; 2002WO-US15125.

XX PR 14-MAY-2001; 2001US-290907P.

XX PA (JIHH/) JI H H.

XX PI Ji HH;

XX DR WPI; 2003-129281/12.

XX PS Isolating a nucleic acid encoding a protein that binds with a target
 PT antigen for amplifying and cloning sequences containing a sequence
 PT encoding a variable domain sequence by employing a transcription-based
 PT amplification scheme -

XX SQ Disclosure; Column 24; 46pp; English.

XX The invention relates to a method for isolating a nucleic acid encoding
 CC a protein that binds with a target antigen for amplifying and cloning
 CC sequences containing a sequence encoding a variable domain sequence by
 CC employing a transcription-based amplification scheme. The method is
 CC useful for amplifying and cloning sequences containing a sequence
 CC encoding a variable domain sequence. The present sequence is a linker
 CC peptide used to illustrate the method of the invention.

XX SQ Sequence 5 AA;

Query Match 100.0%; Score 5; DB 24; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
 Db 1 GGGGS 5

RESULT 65

ABP72514
 ID ABP72514 standard; Peptide; 5 AA.

XX AC ABP72514;

XX DT 23-MAY-2003 (first entry)

XX DE Peptide linker.

XX KW Glycoprotein VI; GPVI; thrombolytic; anticoagulant; antiangiinal;
 KW cardiovascular; linker; drug screening; nonadhesive;
 XX surface coating.

XX OS Synthetic.

XX PN WO2003008454-A2.

XX PD 30-JAN-2003.

XX PF 12-JUL-2002; 2002WO-EP07796.

XX PR 18-JUL-2001; 2001EP-0116717.

XX PA (MERE) MERCK PATENT GMBH.

XX PI Burger C, Gleitz J, Frech M;

XX DR WPI; 2003-229561/22.

XX New Glycoprotein VI (GPVI) fusion protein for preventing or treating
 PT thrombotic or cardiovascular disorders, comprises a tag molecule and a
 PT non-immunoglobulin molecule, e.g. protein or oligopeptide with
 PT biological activity of GPVI -

XX PS Disclosure; Page 11; 42pp; English.

XX The present sequence is that of a linker peptide suitable for
 CC use in fusion proteins of the invention. A multimer (up to a
 CC pentamer) of the peptide is preferably used. The fusion proteins
 CC comprise glycoprotein VI (GPVI) linked to a tag such as myc, GST,
 CC HA, FLAG, STREP or, preferably, the Fc portion of an immunoglobulin,
 CC and may also include a signal sequence for protein secretion. DNA
 CC encoding the fusion protein can be incorporated into an expression
 CC vector and secreted fusion protein purified from eukaryotic host cells.
 CC The fusion protein, especially Fc-GPVI and GPVI-Fc (see ABP72518-19)
 CC can be used to screen for agonists or antagonists of GPVI-collagen
 CC and/or platelet-collagen interactions, and for the treatment of
 CC thrombotic and cardiovascular events and disorders related to
 CC GPVI-collagen and/or platelet-collagen interactions including
 CC increased platelet activation with collagen, atherosclerotic plaque
 CC rupture, unstable angina or during surgical treatment such as
 CC percutaneous transluminal coronary angioplasty (claimed). The
 CC fusion proteins are also useful for coating artificial surfaces to
 CC render them nonadhesive for cells, for modifying intraocular lenses
 CC to lessen the thrombogenicity of the lens material, for contacting
 CC the lens surface, and for covalent crosslinking to modify the lens
 CC material (all claimed).

XX SQ Sequence 5 AA;

Query Match 100.0%; Score 5; DB 24; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5

Db 1 GGGGS 5

RESULT 66

BG75993	ABG75993 standard; Peptide; 5 AA.	ID	ABG74221 standard; peptide; 5 AA.
D	X	AC	ABG74221;
C	X	XX	
X	ABG75993;	D7	
T	30-APR-2003 (first entry)	XX	16-APR-2003 (first entry)
X		DE	Zinc finger protein fusion proteins, peptide linker #5.
X	Spacer peptide #1.	XX	
E	Green fluorescent protein; GFP; FRET; spacer peptide;	KW	Zinc finger protein; vascular endothelial growth factor; VEGF; ZFP;
W	fluorescent resonance energy transfer; tandem fluorescent protein;	KW	VP16 activation domain; VEGF3a/1; VEGF1; SV-40; NLS; cancer;
W	fluorescent resonance energy transfer; tandem fluorescent protein;	KW	nuclear localisation sequence; genetic disease; ischaemia; arthritis;
W	enzymatic assay; Alzheimer's disease; hypertension; inflammation;	KW	vascular disease; immunological disorder; infection; fruit ripening;
W	apoptosis; AIDS; acquired immunodeficiency syndrome.	KW	plant disease resistance; sugar composition; oil composition; HIV;
X	Synthetic.	KW	plant yield; colour; diabetic retinopathy; macular degeneration;
S		KW	rheumatoid arthritis; psoriasis; human immunodeficiency virus infection;
N	US2002164674-A1.	KW	sickle cell anaemia; Alzheimer's disease; muscular dystrophy;
N		KW	neurodegenerative disease; vascular disease; cystic fibrosis; stroke;
X	07-NOV-2002.	XX	plant nutritional characteristic; storage capability.
D		OS	Synthetic.
X	25-JAN-2002; 2002US-0057505.	XX	
F	31-JAN-1997; 97US-0792553.	PN	US2002160940-A1.
R	13-SEP-1999; 99US-0396003.	XX	
R	31-JAN-1996; 96US-0594575.	PD	31-OCT-2002.
X	(REGC) UNIV CALIFORNIA.	XX	
A		PF	28-AUG-2001; 2001US-0942087.
I	Tsien RY, Heim R, Cubitt A;	XX	
I	WPI; 2003-247255/24.	PR	12-JAN-1999; 99US-0229037.
R		XX	(CASE/) CASE C C.
X	New tandem fluorescent protein construct comprising a donor or acceptor	PA	(WOLF/) WOLFFE A.
R	fluorescent protein moiety or a linker moiety that couples the donor	PA	(URNO/) URNOV F.
X	and acceptor moieties, useful in enzymatic assays .	PA	(LAIA/) LAI A.
T		PA	(SNOW/) SNOWDEN A.
T		PA	(TANS/) TAN S.
S	Disclosure; Page 8; 34pp; English.	PA	(GREG/) GREGORY P.
X		PI	Case CC, Wolffe A, Urnov F, Lai A, Snowden A, Tan S, Gregory P;
C	The invention relates to a tandem fluorescent protein construct	XX	WPI; 2003-209226/20.
C	comprising a donor or acceptor fluorescent protein moiety (e.g.	DR	
C	jellyfish green fluorescent protein, GFP) or a linker moiety that	XX	
C	couples the donor and acceptor moieties. Also include are a recombinant	PT	Modulating the expression of an endogenous gene in a cell, e.g. for
C	nucleic acid coding for expression of the tandem fluorescent protein	PT	treating genetic diseases and developing plants with altered
C	construct, an expression vector comprising expression control sequences	PT	phenotypes, comprises contacting a target site in the gene with a zinc
C	operatively linked to a sequence coding for the expression of the	PT	finger protein .
C	tandem fluorescent protein construct, a host cell transfected with the	XX	
C	expression vector, determining whether a sample contains an enzyme or	XX	Disclosure; Page 18; 51pp; English.
C	whether a compound alters the activity of an enzyme, determining the	XX	
C	amount of activity of an enzyme in a cell and testing for cleavage enzyme	CC	The invention relates to modulating the expression of an endogenous
C	activity. The tandem fluorescent protein construct is useful in enzymatic	CC	cellular gene in a cell, comprising contacting a target site in the gene
C	assays, using the principle of fluorescent resonance energy transfer	CC	with a designed or selected zinc finger protein (ZFP), where the ZFP
C	(FRET) between the donor and acceptor moieties. The tandem	CC	comprises a functional domain, and thus modulating the expression of the
C	fluorescent proteins are particularly useful in assays for protease	CC	cellular gene. Effector plasmids expressing vascular endothelial growth
C	activity. Proteases play an essential role in many disease processes e.g.	CC	factors (VEGF) ZFPs fused to the VP16 activation domain were introduced
C	Alzheimer's disease, hypertension, inflammation, apoptosis and AIDS	CC	into cells and were shown to upregulate the VEGF gene. Eukaryotic
C	(acquired immunodeficiency syndrome). Tandem fluorescent proteins	CC	expression vectors were constructed that fuse the VEGF3a/1 and the VEGF1
C	were constructed comprising mutants of GFP with altered fluorescent	CC	ZFPs to the simian virus (SV)-40 nuclear localization sequence (NLS) and
C	spectra. The present sequence represents a flexible spacer peptide	CC	VP16. The method is useful in treating genetic diseases, cancers,
C	which may be incorporated into the tandem fluorescent protein construct.	CC	ischaemia, vascular disease, arthritis, immunological disorders, and
X		CC	fungal, protozoal, bacterial and viral infections. The method is also
X		CC	useful in functional genomics assays, and for developing plants with
X		CC	altered phenotypes, including disease resistance, fruit ripening, sugar
X		CC	and oil composition, yield and colour. The method is also useful for
X		CC	treating diabetic retinopathy, macular degeneration, rheumatoid
X		CC	arthritis, psoriasis, human immunodeficiency virus (HIV) infection,
X		CC	sickle cell anaemia, Alzheimer's disease, muscular dystrophy, and
X		CC	neurodegenerative diseases, vascular disease, cystic fibrosis, and stroke.
X		CC	ZFP proteins are administered to engineer plants for traits including
X		CC	increased disease resistance, modification of structural and storage
X		CC	polysaccharides, flavours, proteins and fatty acids, fruit ripening,
X		CC	yield, colour, nutritional characteristics and improved storage
X		CC	capability. The present sequence is a peptide linker suitable for
X		CC	use in linking the ZFP to a protein or domain of interest.

Sequence 5 AA;
 Query Match 100.0%; Score 5; DB 24; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGGGS 5
 |||||
 b 1 GGGGS 5

RESULT 67
 BG74221

XX SQ Sequence 5 AA;
 Query Match 100.0%; Score 5; DB 24; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGGS 5
 DB 1 GGGGS 5

RESULT 68
 ABP71088
 ID ABP71088 standard; peptide; 5 AA.
 AC ABP71088;
 DT 14-APR-2003 (first entry)
 DE Amino acid sequence of a peptide linker.
 KW HAM; homologue of attractin/mahogany; immunosuppressive; cytostatic;
 KW antiinflammatory; cardiant; osteopathic; gene therapy; linker.
 OS Synthetic.
 XX WO200297120-A1.
 XX 05-DEC-2002.
 XX 23-MAY-2002; 2002WO-US16391.
 XX 25-MAY-2001; 2001US-293608P.
 PR 24-SEP-2001; 2001US-324626P.
 XX (IMMV) IMMUNEX CORP.
 XX Anderson DX;
 XX WPI; 2003-140486/13.
 XX New Homologue of Attractin/Mahogany (HAM) polypeptide, useful for
 PT treating HAM-associated disorder consisting of inflammatory,
 PT autoimmune, cell proliferative or cardiovascular disorders -
 XX Disclosure; Page 14; 89pp; English.
 XX The invention relates to Homologue of Attractin/Mahogany (HAM)
 CC polypeptides and encoding polynucleotides. The HAM polypeptides can be
 CC expressed by standard recombinant methodology. The HAM polypeptides are
 CC useful for treating HAM-associated disorder consisting of inflammatory,
 CC autoimmune, graft-versus-host, neurological, myelination, cell
 CC proliferative, cardiovascular, haematologic, liver, metabolic, weight or
 CC bone disorder. The present sequence represents a peptide linker.
 XX Sequence 5 AA;
 Query Match 100.0%; Score 5; DB 24; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGGS 5
 DB 1 GGGGS 5

RESULT 69
 AAO26593
 ID AAO26593 standard; Peptide; 5 AA.
 AC AAO26593;
 XX

DT 28-MAR-2003 (first entry)
 XX Fusion protein related peptide #1.
 DE Fungicide; fusion protein; cell surface layer; target cell; specificity;
 KW fungal infection.
 XX Unidentified.
 OS JP2002253245-A.
 PN 10-SEP-2002.
 XX 28-FEB-2001; 2001JP-0055200.
 PF 28-FEB-2001; 2001JP-0055200.
 PR (TOYW) TOYOTA CHUO KENKYUSHO KK.
 PA WPI; 2003-150965/15.
 DR A fusion protein useful for protecting an organism against fungi
 XX comprises a first peptide site acting on cell surface of a target cell
 PT and a second peptide site which interacts with cell surface -
 XX Example 2; Page 5; 15pp; Japanese.
 XX The invention relates to a novel fusion protein containing a first
 CC peptide site which can act on the cell surface layer of a target cell and
 CC a second peptide site which can combine to the constituent of the cell
 CC surface layer of the target cell. The fusion protein is useful for
 CC protecting an organism against fungal infection. The method can be used
 CC for enhancing specificity of a membrane-acting peptide to a specific
 CC target cell. This sequence represents a peptide relating to the fusion
 CC protein of the invention.
 XX Sequence 5 AA;
 Query Match 100.0%; Score 5; DB 24; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGGS 5
 DB 1 GGGGS 5

RESULT 70
 ABB98913
 ID ABB98913 standard; Peptide; 5 AA.
 XX ABB98913;
 AC 28-MAR-2003 (first entry)
 DT Variable region anti-bisphenol A antibody chain-related linker peptide.
 XX Variable region; anti-bisphenol A; antibody; murine; heavy chain;
 KW light chain.
 XX Mus sp.
 OS JP2002253259-A.
 PN 10-SEP-2002.
 PD 02-MAR-2001; 2001JP-0058673.
 PF 02-MAR-2001; 2001JP-0058673.
 PR (BIOS-) BIO APPLIED SYSTEMS KK.
 PA WPI; 2003-096537/09.
 DR

X Gene encoding anti-bisphenol A antibody, a recombinant protein and its
 T preparation, a DNA, a vector, a transformant, preparation of a
 T recombinant protein, a kit for determining bisphenol A -
 X Disclosure; Page 4; 19pp; Japanese.

X The present invention relates to sequences for murine heavy chain
 C variable region or light chain variable region of anti-bisphenol A
 C antibody (AB221157-AB221164 and AB98905-AB98912). The sequences are
 C useful for the preparation of recombinant protein. The present sequence
 C is a linker peptide, which was used in the invention.

X Sequence 5 AA;

Query Match 100.0%; Score 5; DB 24; Length 5;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGGGS 5

b |||||
 1 GGGGS 5

RESULT 71

BP56232

D ABP56232 standard; Peptide; 5 AA.

C ABP56232;

X 28-MAR-2003 (first entry)

X Single chain antibody peptide spacer SEQ ID NO:6.

X Mouse; prion protein dimer; fusion protein; vaccination; vaccine;
 W prion protein; virucide; immunostimulant; spongiform encephalopathy;
 W transmissible spongiform encephalopathy.

X Synthetic.

N EP1251138-A1.

X 23-OCT-2002.

X 19-APR-2001; 2001EP-0109707.

X 19-APR-2001; 2001EP-0109707.

X (SCHA/) SCHAETZL H.

A Schaetzl H;

R WPI; 2003-095100/09.

X New prion protein dimer, useful as a vaccine against spongiform
 T encephalopathy -

S Disclosure; Page 3; 23pp; English.

X The present invention describes a prion protein which is a homodimer or
 C heterodimer (1). Also described: (1) an antibody which specifically
 C binds the protein; (2) a DNA sequence encoding the protein; and (3) an
 C expression vector containing the DNA. (1) has virucide and
 C immunostimulant activities. The protein can be used for producing
 C antibodies, and the protein, antibody, DNA and vector are useful for
 C preparing a vaccine for preventing or treating a transmissible
 C spongiform encephalopathy. There is no current vaccine for spongiform
 C encephalopathies. The present sequence represents a peptide spacer
 C which is given in the exemplification of the present invention.

X Sequence 5 AA;

Query Match 100.0%; Score 5; DB 24; Length 5;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGGS 5

Db |||||
 1 GGGGS 5

RESULT 72

AAE33354

ID AAE33354 standard; peptide; 5 AA.

XX AAE33354;

DT 02-APR-2003 (first entry)

DE Linker peptide used in the invention.

XX MUC-1; 3D; variable light domain; VL; variable heavy domain; VH;

KW diabody; cancer; antibody; therapy.

OS Synthetic.

XX WO200279429-A2.

PN 10-OCT-2002.

XX 28-MAR-2002; 2002WO-US09735.

PF 30-MAR-2001; 2001US-280721P.

XX (REGC) UNIV CALIFORNIA.

XX Denardo SJ, Winthrop MD, Denardo GL;

XX WPI; 2003-046804/04.

XX Novel antibody that specifically binds to cancer antigen MUC-1 useful
 PT for detecting a cell bearing MUC-1 antigen, comprises variable light or
 PT variable heavy domains of antibodies 12E, 3D, A5 or C4 -

PS Disclosure; Page 22; 75pp; English.

XX The invention relates to a novel antibody that specifically binds to the
 CC cancer antigen MUC-1. The antibody comprises a domain having a sequence
 CC of a polypeptide selected from 12E variable light (VL) or variable heavy
 CC (VH) domain, 3D VL or VH domain, A5 VL or VH domain and C4 VL or VH
 CC domain. Antibodies of the invention are useful for detecting a cell
 CC bearing a MUC-1 antigen. The invention is useful for producing a variety
 CC of human or humanised antibodies or diabodies. The invention is also
 CC useful for treating cancer. The present sequence is a linker peptide
 CC used in the invention.

XX Sequence 5 AA;

Query Match 100.0%; Score 5; DB 24; Length 5;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGGS 5

Db |||||
 1 GGGGS 5

RESULT 73

ABU55752

ID ABU55752 standard; peptide; 5 AA.

XX ABU55752;

DT 18-MAR-2003 (first entry)

XX Polylinker peptide #4 relating to invention of M. tuberculosis antigens.
 DE

XX
KW Mycobacterium tuberculosis antigenic polypeptide; immune response;
KW tuberculosis infection; polylinker peptide.

OS Synthetic.

PN US6465633-B1.

XX 15-OCT-2002.

PD 23-DEC-1999; 99US-0470191.

XX 24-DEC-1998; 98US-113952P.

XX (CORI-) CORIXA CORP.

PA Skeiky Y;

XX WPI; 2003-147072/14.

XX Novel isolated mycobacterial polynucleotide, useful for treating,
PT preventing or diagnosing Mycobacterium tuberculosis infection, for
PT producing Mycobacterium tuberculosis secretory polypeptides and DNA
PT vaccines

PS Disclosure; Column 91; 48pp; English.

XX The present invention relates to the isolation of polynucleotide
CC sequences encoding Mycobacterium tuberculosis antigenic polypeptides.
CC The polynucleotide sequences of the invention are useful for treating,
CC preventing, and diagnosing M. tuberculosis infection, for producing
CC M. tuberculosis secretory polypeptides, for producing DNA vaccines,
CC for diagnostic purposes, as molecular probes or primers to detect
CC the presence of bacteria in a biological sample, for inducing and/or
CC enhancing immune responses to M. tuberculosis, and in gene therapy.
CC ABUS5749-ABUS5754 represent flexible polylinker peptides.
CC Note: The present sequence is given in the Sequence listing but is
CC not mentioned elsewhere in the specification.

XX Sequence 5 AA;

Query Match 100.0%; Score 5; DB 24; Length 5;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5

Db 1 GGGGS 5

RESULT 74

ABG72271

ID ABG72271 standard; peptide; 5 AA.

XX AC ABG72271;

XX 06-MAR-2003 (first entry)

XX Linker peptide #2 used in recombinant immunotoxin, scFv(UCHT-1)-PE38.

XX Recombinant immunotoxin; scFv(UCHT-1)-PE38; single chain Fv fragment;
KW murine; anti-human CD3-monoclonal antibody binding domain; UCHT-1;
KW Pseudomonas aeruginosa; exotoxin A; ADP-ribosylating; translocation;
KW T-cell mediated disorder; organ transplantation rejection;
KW autoimmune disease; graft versus host disease; bone marrow transplant;
KW acquired immunodeficiency syndrome; AIDS; T-cell leukaemia;
KW T-cell lymphoma; T-cell-mediated autoimmune disease; type I diabetes;
KW systemic lupus erythematosus; rheumatoid arthritis; myasthenia gravis;
KW multiple sclerosis; immunological tolerance; osteoporosis;
KW aplastic anaemia; Gaucher's disease; thalassaemia.

XX Synthetic.

OS

PN US2002142000-A1.

XX 03-OCT-2002.

XX 10-JAN-2000; 2000US-0480236.

XX 10-JAN-2000; 2000US-0480236.

XX (DIGA/) DIGAN M E.

PA (LAKE/) LAKE P.

PA (WRIGHT/) WRIGHT R M.

XX Digan ME, Lake P, Wright RM;

XX WPI; 2003-155935/15.

XX Novel recombinant immunotoxin polypeptide useful for treatment or
PT prophylaxis of T-cell mediated disorders or organ transplantation
PT rejection, comprises a CD3-binding domain and a Pseudomonas exotoxin
PT mutant

XX Disclosure; Page 8; 58pp; English.

XX The present invention relates to a novel recombinant immunotoxin,
CC scFv(UCHT-1)-PE38, and the polynucleotide sequence encoding it. The
CC recombinant immunotoxin comprising a single chain (sc) Fv fragment
CC of murine anti-human CD3-monoclonal antibody (UCHT-1) binding domain
CC fused to a Pseudomonas aeruginosa exotoxin A (PE38) mutant having
CC ADP-ribosylating and translocation functions but substantially
CC diminished cell-binding ability. The recombinant immunotoxin is useful
CC for the treatment or prophylaxis of T-cell mediated disorders, organ
CC transplantation rejection or autoimmune disease in a patient. For
CC example, scFv(UCHT-1)-PE38 may be used for treatment or prophylaxis
CC against graft versus host disease in a patient undergoing a
CC bone marrow transplant, for the treatment or prophylaxis of transplant
CC rejection in a patient to undergo a bone marrow transplant, or for
CC conditioning a patient to be transplanted with cells, tissue or
CC organ of a donor. scFv(UCHT-1)-PE38 is also useful for treating
CC acquired immunodeficiency syndrome (AIDS), T-cell leukaemias or
CC lymphomas, T-cell-mediated autoimmune disease such as systemic lupus
CC erythematosus, type I diabetes, rheumatoid arthritis, myasthenia
CC gravis, multiple sclerosis, for inducing immunological tolerance, and
CC for treating diseases curable or treatable by bone marrow
CC transplantation, including leukaemias, osteoporosis, aplastic anaemia,
CC Gaucher's disease, and thalassaemia. The present sequence represents
CC a linker peptide used in the construction of recombinant immunotoxin,
CC scFv(UCHT-1)-PE38.

XX Sequence 5 AA;

Query Match 100.0%; Score 5; DB 24; Length 5;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5

Db 1 GGGGS 5

RESULT 75

AAE30437

ID AAE30437 standard; peptide; 5 AA.

XX AC AAE30437;

XX 24-FEB-2003 (first entry)

XX Peptide linker #2 used to link zinc finger proteins.

XX DNA binding specificity; zinc finger protein; specificity optimisation;
KW ZFP; gene expression.

XX Unidentified.

OS

WO200277227-A2.
03-OCT-2002.
20-NOV-2001; 2001WO-US43568.
20-NOV-2000; 2000US-0716637.
(SANG-) SANGAMO BIOSCIENCES INC.
Eisenberg SP, Liu Q, Jamieson A, Rebar E;
WPI; 2003-029936/02.
Enhancing the binding specificity of a zinc finger protein, comprises substituting amino acids in the protein for residues in the target sequence to make a modified binding protein -
Disclosure; Page 19; 55pp; English.
The present invention relates to a novel method of enhancing the binding specificity of a binding protein (zinc finger protein (ZFP)). The method involves substituting one or more amino acids at positions in the binding protein that affect the specificity of the binding protein for residues in the target sequence to make a modified binding protein. The method is useful in optimising the specificity of a binding protein, in modulating the expression of a target gene in a subject and in diagnostic methods for sequence-specific detection of a target nucleic acid in a sample. The present sequence is a peptide linker used to link zinc finger proteins.

Query Match 100.0%; Score 5; DB 24; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGGGS 5
b 1 GGGGS 5
|||||

RESULT 76
AAE30955 standard; peptide; 5 AA.
AAE30955;
24-FEB-2003 (first entry)
Glycine rich peptide linker used to construct GLP-1 fusion protein.
Glucagon-like peptide 1; GLP-1; albumin; immunoglobulin; Ig; therapy; non-insulin diabetes mellitus; obesity; antidiabetic; anorectic.
Synthetic.
WO200246227-A2.
13-JUN-2002.
29-NOV-2001; 2001WO-US43165.
07-DEC-2000; 2000US-251954P.
(ELIL) LILLY & CO ELI.
Glaesner W, Micanovic R, Tschang SR;
WPI; 2003-018534/01.

Novel heterologous fusion protein, useful for treating non-insulin dependent diabetes mellitus or obesity, comprises a glucagon-like peptide 1 compound fused to human albumin or to the Fc portion of an immunoglobulin -
Claim 3; Page 94; 200pp; English.
The invention relates to a heterologous fusion protein comprising a first polypeptide fused to a second polypeptide, where the polypeptides has a N-terminus and a C-terminus and the first polypeptide is a glucagon-like peptide 1 (GLP-1) compound and the second is a human albumin or its analogue or fragment, or the Fc portion of an immunoglobulin (Ig) or its analogue or fragment, where the C-terminus of first polypeptide is fused to the N-terminus of the second polypeptide. The invention is useful for normalising blood glucose levels in mammal, for treating a patient with non-insulin diabetes mellitus or obesity, or for the manufacture of CC medicament for treating the above mentioned diseases. The present CC sequence is glycine rich peptide linker used to construct GLP-1 fusion protein.
Sequence 5 AA;
Query Match 100.0%; Score 5; DB 24; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 1 GGGGS 5
|||||

RESULT 77
AAO16029 standard; peptide; 5 AA.
AAO16029;
20-FEB-2003 (first entry)
Cyclic single-strand anti-tumour antibody-related peptide #2.
Cyclic single-strand anti-tumour antibody; anti-tumour Fab; anti-human Cd3 Fab; anti-human CD28 Fab; immunotherapy; tumour; ovarian cancer.
Unidentified.
WO200283738-A1.
24-OCT-2002.
10-APR-2002; 2002WO-CN00252.
11-APR-2001; 2001CN-0110554.
(GENE-) INST GENETICS CAS.
(BEIJ-) BEIJING ABT GENETIC ENG TECHNOLOGY CO LT.
Huang H, Cheng J, Wang X, Song L, Zhang Z, Lin Q, Gu Y;
WPI; 2003-075525/07.
N-PSDB; AAL51055.
Cyclic single-strand tri-specific antibody made from antitumor Fab, single-strand antibody and modified anti-human CD3 or CD28 Fab, applicable in early-stage diagnosis, and prevention or treatment of tumors e.g. ovarian cancer -
Claim 6; Fig 4A; 49pp; Chinese.
The invention comprises a cyclic single-strand anti-tumour antibody, the antibody of the invention is obtained by combining anti-tumour Fab and modified anti-human Cd3 or CD28 Fab. The antibody is useful in the

CC immunotherapy of tumours - particularly ovarian cancer, including the
 CC sensitive early-stage diagnosis and post-operative eradication of
 CC disease. The present amino acid sequence represents a peptide which was
 CC used in the invention.

XX
 XX Sequence 5 AA;
 SQ
 Query Match 100.0%; Score 5; DB 24; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
 Db 1 GGGGS 5

RESULT 78
 AAR62168
 ID AAR62168 standard; Protein; 6 AA.

XX

AC AAR62168;

XX 25-MAR-2003 (updated)

DT 03-MAY-1995 (first entry)

XX

DE U1 snRNP 70K protein amino acids 5-10, homologous to HSV-1 IE motif.

XX

XX Small ribonucleoprotein complex; U1 snRNP; 70K protein; epitope;
 KW autoantibody; immunoinfective cluster virus; nuclear protein antigen;
 KW systemic rheumatic disorder; herpes simplex virus; HSV-1;
 KW immediate early protein; systemic lupus erythematosus; scleroderma.

XX

OS Homo sapiens.

XX

PN WO9420141-A1.

XX

PD 15-SEP-1994.

XX

PF 10-MAR-1994; 94WO-US02631.

XX

PR 11-MAR-1993; 93US-0029850.

XX

PA (UYSC-) UNIV SOUTHERN CALIFORNIA.

XX

PI Douvas A, Ehresmann G, Takehana Y;

XX

DR WPI; 1994-302689/37.

XX

XX Methods for treating immunoinfective cluster virus infections -
 PT utilise antibodies or fragments characteristic of auto antibodies
 PT produced by patients with rheumatic disorders

XX

PS Disclosure; Page 63; 106pp; English.

XX

XX A comparison of the U1 snRNP 70K protein sequence with proteins
 CC from immunoinfective cluster viruses revealed widespread
 CC homologues. The importance of these homologous motifs is that they
 CC are epitopes for autoantibodies occurring in high titres in systemic
 CC rheumatic disorders. Sera from such patients could be used for
 CC treatment of immunoinfective cluster virus infections.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX

SQ Sequence 6 AA;

Query Match 100.0%; Score 5; DB 15; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
 Db 2 GGGGS 6

RESULT 79

AAW17095

ID AAW17095 standard; peptide; 6 AA.

XX

AC AAW17095;

XX

DT 14-SEP-1999 (first entry)

XX

DE Gly(5)-Ser linker peptide for chimeric protein construct.

XX

XX Haematopoietic protein; human; granulocyte-colony stimulating factor;
 KW G-CSF; interleukin; c-mpl ligand; linker; gene therapy; aplastic anaemia;
 KW stem cell expansion; leukopaenia; neutropaenia; vector; bone marrow;
 KW thrombocytopaenia; blood cell activation; growth.

XX

OS Synthetic.

XX

PN WO9712985-A2.

XX

PD 10-APR-1997.

XX

PF 04-OCT-1996; 96WO-US15774.

XX

PR 05-OCT-1995; 95US-0004834.

XX

PA (SEAR) SEARLE & CO G D.

XX

PI Bauer SC, Baum CM, Caparon MH, Feng Y, Giri JG;

XX

PI Klein BK, Lee SC, McKearn JP, McWhirter CA, Staten NR;

XX

PI Summers NL, Zurluh L;

XX

DR WPI; 1997-226228/20.

XX

XX Multi-functional haematopoietic receptor agonists - used to
 stimulate the production of haematopoietic cells in patients

XX

PS Disclosure; Page 33; 616pp; English.

XX

XX The invention relates to a novel haematopoietic protein (HP) comprising
 CC an amino acid (AA) sequence of formula: R1-L1-R2; R2-L1-R1; R1-R2; or
 CC R2-R1; where R1 and R2 are independently selected from: (I) a modified
 CC human granulocyte-colony stimulating factor (hG-CSF) AA sequence;
 CC (II) a modified human interleukin-3 (hIL-3) AA sequence; (III) a modified
 CC human c-mpl ligand; and a colony stimulating factor (CSF); and L1 = a
 CC linker capable of linking R1 to R2. This sequence represents an example
 CC of a linker used to construct the proteins of the invention.

XX

XX Vectors comprising the nucleic acid molecules are useful for the
 CC recombinant production of HP. The nucleic acid molecules are useful in
 CC gene therapy. The HP's are useful for stimulating the production of
 CC haematopoietic cells in patients, selective ex vivo expansion of stem
 CC cells and for treatment of haematopoietic disorders. Disorders that
 CC can be treated include leukopaenia, neutropaenia, aplastic anaemia and
 CC thrombocytopaenia. In vitro uses include the ability to stimulate bone
 CC marrow and blood cell activation and growth before infusion into the
 CC patients.

XX

SQ Sequence 6 AA;

Query Match 100.0%; Score 5; DB 18; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
 Db 2 GGGGS 6

RESULT 80

AAW21967

ID AAW21967 standard; peptide; 6 AA.

XX

AC AAW21967;

XX

DT 03-DEC-1997 (first entry)

	Linker #1 for immunotoxin containing Pseudomonas exotoxin.	
K	PCR; primer; amplify; polymerase chain reaction; antibody; immunotoxin;	
E	variable heavy chain; VH; murine monoclonal antibody; Lewisy; carcinoma;	
N	carbohydrate antigen; Pseudomonas exotoxin; proteolytic activation;	
M	cytotoxic activity; tumour; autoimmune condition; rheumatoid arthritis;	
W	graft versus host disease; organ transplant rejection; type I diabetes;	
W	multiple sclerosis; systemic lupus erythematosus; myasthenia gravis;	
M	T cell; B cell; cytosol; bone marrow; transplant; therapy.	
S	Synthetic.	
S	WO9713529-A1.	
N	17-APR-1997.	
D	11-OCT-1996; 96WO-US16327.	
F	13-OCT-1995; 95US-0005388.	
R	(USSH) US DEPT HEALTH & HUMAN SERVICES.	
A	Kuan C, Pastan I;	
I	WPI; 1997-235666/21.	
R	Immunotoxin(s) comprising Pseudomonas exotoxin linked to	
X	di-sulphide stabilised variable heavy and light chain regions of an	
T	antibody - useful for killing target cells bearing characteristic	
T	marker	
T	Claim 5; Page 49; 64pp; English.	
X	AAW21967-W21969 represent linkers used in the immunotoxins of the	
S	invention. The immunotoxins bind to target cells, and comprise, a	
C	Pseudomonas exotoxin (PE) that does not need proteolytic activation for	
C	cytotoxic activity fused to a VH framework region of an Fv antibody (Ab)	
C	fragment. The VH chain region is bound through at least one disulphide	
C	bond to a variable light (VL) chain framework region. The PE is lacking	
C	residues 1-279 and is at least 10-fold more cytotoxic to the target	
C	cells than an immunotoxin comprising PE attached to a VH chain framework	
C	region of an Fv Ab fragment lacking a disulphide bond to a VL chain	
C	framework region. These sequences are used to join the VH chain region to	
C	the PE. The immunotoxins can be used for killing target cells in the	
C	treatment of tumours, autoimmune conditions, graft versus host disease,	
C	organ transplant rejection, type I diabetes, multiple sclerosis,	
C	rheumatoid arthritis, systemic lupus erythematosus, myasthenia gravis,	
C	etc., all caused by T and B cells. They can also be used to deliver an	
C	antibody to the cytosol of a cell, and in vitro in the elimination of	
C	harmful cells from bone marrow before transplant. The immunotoxins have	
C	high cytotoxicity to target cells and a small size to provide greater	
C	penetration to target cells.	
C	Sequence 6 AA;	
C	Query Match 100.0%; Score 5; DB 18; Length 6;	
C	Best Local Similarity 100.0%; Pred.No. 9.3e+05;	
C	Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
/	1 GGGS 5	
/		
/	2 GGGS 6	
SULT 81		
G78084		
ABG78084 standard; Peptide; 6 AA.		
ABG78084;		
15-NOV-2002 (first entry)		

CC their modulators are also useful for treating pancreatic disorders,
 CC heart disorders, pituitary-related disorders, disorders of the adrenal
 CC cortex, thyroid, placenta, gastric disorders, pulmonary disorders,
 CC disorders of the skeletal muscle, hepatic disorders (e.g. jaundice,
 CC hepatitis), renal disorders (e.g. glomerulonephritis), testicular
 CC disorders, uterine disorders, disorders of the brain, central nervous
 CC system disorders including Alzheimer's disease, Parkinson's disease,
 CC multiple sclerosis, brain cancers, eye disorders (e.g. glaucoma,
 CC cataract), prostate and ovarian disorders, intestinal disorders,
 CC colonic, and leukocytic disorders. ABG78024-ABG78125 represent ITALY,
 CC Lox-2, STRIFE, TRASH, BBSF, LRSG and STMSF amino acid sequences and
 CC related peptides of the invention.

SQ Sequence 6 AA;
 Query Match 100.0%; Score 5; DB 23; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
 Db 1 GGGGS 5
 |||||

RESULT 82
 ABB84094
 ID ABB84094 standard; peptide; 6 AA.

XX ABB84094;
 AC ABB84094;
 DT 09-SEP-2002 (first entry)
 DE Human scDb CEAGal PA & PB linker peptide (wild type).
 KW Antibody; multimer; cytostatic; detection; immunoassay; tumour.

OS Homo sapiens.
 XX DE10060140-A1.
 XX 06-JUN-2002.

XX 04-DEC-2000; 2000DE-1060140.
 XX 04-DEC-2000; 2000DE-1060140.
 XX (VECT-) VECTRON THERAPEUTICS INT AG.
 XX Kontermann R;

XX WPI; 2002-520984/56.
 XX New multimer of single-chain antibodies, useful e.g. for diagnosis or
 PT drug delivery, has four variable chains linked through peptides of
 PT controlled length -
 XX Example 1; Page 10; 16pp; German.

XX This invention describes a novel multimer containing at least two
 CC antibodies, each, independently, comprising a polypeptide chain of
 CC structure VI-PA-V2-PM-V3-PB-V4 where VI-V4 = variable domains; PA, PB
 CC and PM = peptide linkers with PA and PB containing 0 or 1 amino acid. The
 CC products of the invention have cytostatic activity. The multimers
 CC described can be used for diagnostic detection, in essentially standard
 CC immunoassay methods, and for binding at least one component and/or fused
 CC protein or peptide to a cell, especially to deliver a toxin or for
 CC infection, transformation or transfection of the cell. (I), or nucleic
 CC acid (II) encoding it, or vectors containing (II), are useful for
 CC treating tumours. The multimers are relatively small (about 110 kDa for a
 CC dimer) but have multiple binding sites to ensure high binding affinity.
 CC This sequence represents a wild-type PA and PB type linker peptide
 CC described in the method of the invention.

SQ Sequence 6 AA;
 Query Match 100.0%; Score 5; DB 23; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
 Db 2 GGGGS 6
 |||||

RESULT 83
 AAO14388
 ID AAO14388 standard; Peptide; 6 AA.

XX AAO14388;
 AC AAO14388;
 XX 24-MAY-2002 (first entry)
 DT 24-MAY-2002 (first entry)
 XX Linker peptide for use in a fusion protein #2.

XX Human: metalloproteinase-like/disintegrin-like protein; MPD;
 KW angiogenesis; endothelial cell migration; integrin ligand binding;
 KW ocular disorder; cancer; inflammation; osteoporosis; restenosis;
 KW cytostatic; antiinflammatory; osteopathic; vasotropic; thrombolytic;
 KW vulnery; anti allergic; antibacterial; virucide; protozoacide; cardiant;
 KW antianaemic; analgesic; thrombosis; tissue repair; neural disorder;
 KW infection; linker.

XX Synthetic.
 XX Key Location/Qualifiers
 FH Misc-difference 6 /label= Xaa
 FT /note= "Xaa=one or more repeats of the present sequence"

XX WO200210406-A2.
 XX 07-FEB-2002.
 XX 27-JUL-2001; 2001WO-US23734.
 XX 28-JUL-2000; 2000US-221838P.

XX (IMKV) IMMUNEX CORP.
 XX Dubose RE, Wiley SR, Black RA;
 XX WPI; 2002-195965/25.
 XX Novel purified human metalloproteinase-disintegrin polypeptide, useful
 PT for treating ocular and inflammatory disorders, osteoporosis, cancer,
 PT restenosis, thrombosis, and chronic pain conditions -
 XX Disclosure; Page 20; 101pp; English.

XX The present invention provides a number of human metalloproteinase-like
 CC /disintegrin-like (MPD) proteins and their fragments. These can be used
 CC to identify treatments for ocular disorders, malignant and metastatic
 CC conditions, inflammatory diseases, osteoporosis and other conditions
 CC mediated by accelerated bone resorption, restenosis, inappropriate
 CC platelet activation, recruitment or aggregation, thrombosis, conditions
 CC requiring tissue repair or wound healing, endothelial migration,
 CC angiogenesis, allergies, reproductive, neurological and vascular
 CC conditions, infections, cardiovascular disorders such as myocardial
 CC infarction, chronic pain conditions, endocrine system disorders,
 CC gastrointestinal system disorders, genitourinary system disorders,
 CC anaemia and haematological disorders. The present sequence is a linker
 CC peptide described in the exemplification of the invention.

SQ Sequence 6 AA;
 Query Match 100.0%; Score 5; DB 23; Length 6;

CC (e.g. arthritis, Graves disease, multiple sclerosis (MS), autoimmune
 CC myocarditis, diabetes and lupus) by administering a peptide construct
 CC comprising a TCEL peptide bonded to an antigenic peptide associated with
 CC the disorder. Unlike prior art peptide conjugates, a modified version of
 CC peptide G has long range stabilisation and also enhances the immune
 CC response. AA082019-AA082114 represent T-cell specific binding ligand
 CC peptides, peptide constructs or peptides used in their construction.
 XX
 SQ Sequence 6 AA;

Query Match 100.0%; Score 5; DB 23; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e-05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
 |||||
 Db 1 GGGGS 5

RESULT 86

ID AAE34832 standard; peptide; 6 AA.

AC AAE34832;

DT 28-MAY-2003 (first entry)

DE Peptide linker #2.

XX Siglec-12; therapy; organ transplant disorder; diabetes mellitus;
 KW haematologic disorder; Fanconi's aplastic anaemia; septic shock; stroke;
 KW Alzheimer's disease; infection; aschma; cachexia; myocardial infarction;
 KW immunosuppressive; thyronimetic; antibacterial; antiparasitic; cardiant;
 KW liver disorder; cerebroprotective; infection; virucide; neuroprotective;
 KW lymphoma; osteoporosis; dermatological; nootropic; rheumatoid arthritis;
 KW dementia; protozoicide; vasotropic; Hashimoto's thyroiditis; cytostatic;
 KW ischaemia; reperfusion injury; systemic lupus erythematosus; psoriasis;
 KW osteopathic.

XX Unidentified.

XX Key Location/Qualifiers

FT Misc-difference 6 /note= "X at this position represents one or more
 FT repeats of GGGGS"

XX WO200296452-A1.

XX 05-DEC-2002.

XX 29-MAY-2002; 2002WO-US16906.

XX 29-MAY-2001; 2001US-294199P.

XX (IMMV) IMMUNEX CORP.

XX Anderson DM, Marken JS;

XX MPI; 2003-140419/13.

XX Novel substantially purified Siglec-12 polypeptide useful for treating
 PT rheumatoid arthritis, systemic lupus erythematosus, Hashimoto's
 PT thyroiditis, asthma, Alzheimer's disease, dementia, cachexia,
 PT myocardial infarction -

XX Disclosure; Column 16; 47pp; English.

XX The invention relates to Siglec-12 polypeptide and polynucleotide
 CC sequences. Siglec-12 is useful for identifying an agent that modulates
 CC its activity. The invention is useful for treating a Siglec-associated
 CC disorder or disease (e.g. rheumatologic disorder, bone marrow or solid
 CC organ transplant disorder, a graft-versus-host disorder, inflammatory
 CC disorder, autoimmune disorder, neurologic disorder, cell proliferative

CC disorder, infection, cardiovascular disorder, haematologic disorder,
 CC liver disorder or a bone disorder). Antibody that specifically binds
 CC to Siglec-12 is useful for treating a tumour that expresses Siglec-12
 CC polypeptide. The invention is also useful for treating systemic lupus
 CC erythematosus, rheumatoid arthritis, dementia, lymphoma, Hashimoto's
 CC thyroiditis, Alzheimer's disease, bacterial, parasitic, protozoal and
 CC viral infections, asthma, cachexia, myocardial infarction, ischaemia/
 CC reperfusion injury, stroke, diabetes mellitus, psoriasis, Fanconi's
 CC aplastic anaemia, septic shock, osteoporosis, etc. The present sequence
 CC is a peptide linker used in the invention.
 XX

SQ Sequence 6 AA;

Query Match 100.0%; Score 5; DB 24; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e-05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
 |||||
 Db 1 GGGGS 5

RESULT 87

ID ABP71089 standard; peptide; 6 AA.

AC ABP71089;

DT 14-APR-2003 (first entry)

XX Amino acid sequence of a peptide linker.

XX HAM; homologue of attractin/mahogany; immunosuppressive; cytostatic;
 KW antiinflammatory; cardiant; osteopathic; gene therapy; linker.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 6 /note= "Xaa is one or more repeats of GGGGS"

XX WO200297120-A1.

XX 05-DEC-2002.

XX 23-MAY-2002; 2002WO-US16391.

XX 25-MAY-2001; 2001US-293608P.

XX 24-SEP-2001; 2001US-324626P.

XX (IMMV) IMMUNEX CORP.

XX Anderson DM;

XX MPI; 2003-140486/13.

XX New Homologue of Attractin/Mahogany (HAM) polypeptide, useful for
 FT treating HAM-associated disorder consisting of inflammatory,
 FT autoimmune, cell proliferative or cardiovascular disorders -

XX Disclosure; Page 14; 89pp; English.

XX The invention relates to Homologue of Attractin/Mahogany (HAM)
 CC polypeptides and encoding polynucleotides. The HAM polypeptides can be
 CC expressed by standard recombinant methodology. The HAM polypeptides are
 CC useful for treating HAM-associated disorder consisting of inflammatory,
 CC autoimmune, graft-versus-host, neurological, myelination, cell
 CC proliferative, cardiovascular, haematologic, liver, metabolic, weight or
 CC bone disorder. The present sequence represents a peptide linker.

SQ Sequence 6 AA;

Query Match 100.0%; Score 5; DB 24; Length 6;

```

Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGGGS 5
b 1 GGGGS 5

RESULT 88
AR9245
D AAR9245 standard; Peptide; 7 AA.
X C AAR9245;
X T 28-NOV-1996 (first entry)
X T (Gly6)Ser linker.
X E Bioactive fusion protein; interleukin-12; IL-12; p35 subunit;
X W p40 subunit; antitumour; cytokine; tumour; melanoma;
X W fibrosarcoma; renal cell carcinoma; immunotherapy; therapy;
X W retrovirus; vector.
X S Synthetic.
X S
X N W09624676-A1.
X D 15-AUG-1996.
X F 07-FEB-1996; 96WO-US01787.
X R 08-FEB-1995; 95US-0385335.
X A (WHED ) WHITEHEAD INST BIOMEDICAL RES.
X I Lieschke GJ, Mulligan RC;
X R WPI; 1996-384448/38.
X R N-PSDB; AAT35197.
X R New DNA encoding fusion protein, esp. contg. IL-12 p35 and p40
X R subunits - for treatment of established tumours or prevention of
X R tumour establishment
X K Claim 2; Page 69; 118pp; English.
X S Peptide linkers (Gly4Ser)2Ser, (Gly4Ser)3Ser, (Gly4Ser)3 and
X C (Gly6)Ser (AAR9242-45) are used to join the subunits of novel
X C dimeric or multimeric fusion proteins. They have been utilised in
X C the prodn. of bioactive interleukin-12 (IL-12) fusion proteins,
X C linking mouse/human IL-12 p35 subunit (see also AAR9246) to
X C mouse/human IL-12 p40 subunit (AAR9247). DNA encoding such constructs
X C can be incorporated into a retroviral vector (see also AAT35198) to
X C allow dimeric IL-12 prodn. in transfected cells. Tumour cells (esp.
X C CMS-5, B16 or renal carcinoma cells) secreting IL-12 dimer can be
X C used to reduce the size of established tumours and/or increase
X C survival time, esp. in cases of melanoma, fibrosarcoma and renal
X C cell carcinoma.
X C
X C Sequence 7 AA;
Query Match 100.0%; Score 5; DB 17; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGGGS 5
b 3 GGGGS 7

RESULT 89
AY23703
D AAY23703 standard; Peptide; 7 AA.

```

```

XX AAY23703;
XX AC
XX DT 07-SEP-1999 (first entry)
XX XX
XX DE Peptide identified using the method of the invention.
XX XX
XX KW Polysome library; nascent peptide display;
XX KW affinity interaction screening; multiple binding specificity;
XX KW single chain antibody.
XX OS Synthetic.
XX XX US5922545-A.
XX PD 13-JUL-1999.
XX PF 29-JUL-1997; 97US-0902623.
XX XX 17-JAN-1996; 96US-0586176.
XX PR 29-OCT-1993; 93US-0144775.
XX PR 02-SEP-1994; 94US-0300262.
XX PR 25-OCT-1994; 94WO-US12206.
XX PR 29-JUL-1997; 97US-0902623.
XX PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX XX
XX PI Dower WJ, Mattheakis LC;
XX XX WPI; 1999-417975/35.
XX XX
XX PT In vitro peptide and antibody display libraries for identifying
XX PT multiple binding specificities of single chain antibodies
XX XX
XX PS Example 1; Column 71-72; 50pp; English.
XX XX
XX CC The specification describes a method for generating libraries of
XX CC polysomes displaying nascent peptides suitable for affinity
XX CC interaction screening. The method can be used to identify multiple
XX CC binding specificities of single chain antibodies, and comprises
XX CC contacting antigens with a polysome library displaying nascent
XX CC peptides having a single chain antibody segment, separating polysomes
XX CC bound to the antigen species from polysomes not bound to the antigen
XX CC species, and synthesising cDNA from the separated bound polysomes to
XX CC identify single chain antibodies which bind to at least one of the
XX CC antigen species present. The method for generating libraries of
XX CC polysomes displaying nascent peptides is useful for affinity interaction
XX CC screening. The present sequence was identified using the method of the
XX CC invention.
XX SQ Sequence 7 AA;
Query Match 100.0%; Score 5; DB 20; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 3 GGGGS 7

RESULT 90
AAY02129
ID AAY02129 standard; Protein; 7 AA.
XX XX
XX AC AAY02129;
XX XX
XX DT 16-JUL-1999 (first entry)
XX DE Peptide linker used to make multifunctional proteins.
XX XX
XX KW Angiostatin; endostatin; interferon; thrombospondin;
XX KW interferon-inducible protein; platelet factor 4; anti-angiogenic;

```

KW anti-tumor; multifunctional protein; angiogenic-mediated disease;
 KW cancer; diabetic retinopathy; macular degeneration; arthritis;
 KW tumor cell production; peptide linker.

XX Homo sapiens.

XX WO9916889-A1.

XX 08-APR-1999.

PF 30-SEP-1998; 98WO-US20464.

PR 01-OCT-1997; 97US-0060609.

XX (SEAR) SEARLE & CO G D.

PI Bolanowski MA, Caparon MH, Casperson GF, Gregory SA;

PI Klein BK, McKearn JP;

DR WPI; 1999-255098/21.

XX New multifunctional proteins useful for treating angiogenic-mediated diseases

PS Disclosure; Page 112; 121pp; English.

XX The specification describes multifunctional proteins which comprise combinations of angiotensin, endostatin, interferon, thrombospondin, interferon-inducible protein and platelet factor 4, and have anti-angiogenic and/or anti-tumor activity. The multifunctional protein may exhibit useful properties such as having similar or greater biological activity when compared to a single factor or by having improved half-life or decreased adverse side effects, or a combination of these properties. The proteins can be used for treating an angiogenic-mediated disease, e.g. cancer, diabetic retinopathy, macular degeneration, or arthritis. They can also be used for inhibiting the production of tumor cells (characteristic of lung, breast, ovarian, prostate, pancreatic, gastric, colon, renal, bladder cancers; melanoma, hepatoma, sarcoma and lymphoma) in a patient and for inhibiting tumor growth. AA0125-32 represent peptide linkers used to make the multifunctional proteins of the invention.

SQ Sequence 7 AA;

Query Match 100.0%; Score 5; DB 20; Length 7;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5

Db 3 GGGGS 7

RESULT 91

ABB56487

ID ABB56487 standard; Peptide; 7 AA.

XX ABB56487;

DT 25-FEB-2002 (first entry)

DE Human single chain MHC class II molecule linker #11.

XX Human; MHC; major histocompatibility complex; MHC class II; multimer; single chain; immunosuppressive; antidiabetic; antiinflammatory; antianemic; antirheumatoid; antiarthritic; neuroprotective; vaccine; autoimmune disease; insulin dependent diabetes; multiple sclerosis; myasthenia gravis; pernicious anaemia; autoimmune encephalomyelitis; rheumatoid arthritis; systemic lupus erythematosus; linker.

XX Homo sapiens.

OS WO200170245-A1.

XX

PN

XX

PD

XX

PF

XX

PR

XX

PR

XX

PR

XX

PA

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

27-SEP-2001.

22-MAR-2001; 2001WO-US09616.

22-MAR-2000; 2000US-191274P.

15-MAY-2000; 2000US-204249P.

23-JAN-2001; 2001US-264003P.

(CORI-) CORIXA CORP.

Carter D, Zhu S, Arimilli S, Wang A;

WPI; 2001-616371/71.

Multimeric complex for treating autoimmune diseases, comprises first

and second single chain MHC class II molecules, each comprising alpha

and beta domain linked through amino acid linker and multimerization

domain -

Claim 24; Page 131; 147pp; English.

The invention relates to a multimeric complex comprising a first

recombinant single chain major histocompatibility complex (MHC) class

II molecule and a second recombinant single chain MHC class II molecule,

each comprising an alpha domain and a beta domain linked through an

amino acid linker and a multimerization domain. The first and the

second molecule are linked through the multimerization domain to form

a multimeric complex. The complex is useful for treating autoimmune

diseases. It is useful for treating insulin dependent diabetes,

multiple sclerosis, myasthenia gravis, pernicious anaemia, autoimmune

encephalomyelitis (EAE), rheumatoid arthritis and systemic lupus

erythematosus. The present sequence is a linker used to connect the

alpha and beta domains of the single chain MHC class II molecules of

the invention.

Sequence 7 AA;

Query Match 100.0%; Score 5; DB 22; Length 7;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5

Db 3 GGGGS 7

RESULT 92

AAB60900

ID AAB60900 standard; peptide; 7 AA.

XX AAB60900;

XX 03-APR-2001 (first entry)

XX Linker peptide.

XX DTct; diphtheria toxin; insect; baculovirus.

XX Synthetic.

XX US6168932-B1.

XX 02-JAN-2001.

XX 13-JUL-1998; 98US-0114503.

XX 13-JUL-1998; 98US-0114503.

XX (PARK-) PARKER HUGHES INST.

XX Uckun FM, Rostovstev A, Williams MD;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

RESULT 97
 AAE32389
 ID AAE32389 standard; peptide; 7 AA.
 XX
 AC AAE32389;
 XX
 DT 24-MAR-2003 (first entry)
 XX
 DE Linker peptide #1 used in the invention.
 XX
 KW Vascular endothelial growth factor; VEGF; angiogenesis; wound healing;
 KW bone growth; osteoporosis; osteoarthritis; bone reconstruction; ulcer;
 KW lesion; injury; trauma; periodontal condition; protein therapy.
 KW
 DS Unidentified.
 CX
 CN WO200283851-A2.
 CX
 CD 24-OCT-2002.
 CX
 CF 10-APR-2002; 2002WO-US11406.
 CX
 CG 10-APR-2001; 2001US-0832355.
 CX
 CH (GENV-) GENVEC INC.
 CX
 CI Kovesdi I, Kessler PD;
 CX
 CJ WPI; 2003-075536/07.
 CX
 CK New fusion protein comprising a non-heparin-binding vascular
 CK endothelial growth factor (VEGF) peptide portion and a non-VEGF peptide
 CK portion, useful for promoting angiogenesis and/or bone growth in
 CK mammals -
 CK
 CS Disclosure; Page 150; 191pp; English.
 CX
 CC The invention relates to a fusion protein comprising non-heparin binding
 CC vascular endothelial growth factor (VEGF) peptide portion and a non-VEGF
 CC peptide portion useful for promoting angiogenesis and/or bone growth in
 CC mammalian host. The fusion protein is useful for promoting angiogenesis,
 CC wound healing and bone growth. Compositions containing bone growth
 CC promoting fusion protein can be used to treat osteoporosis, rheumatoid
 CC or osteoarthritis, to improve poor bone healing, to promote implant
 CC integration and function of artificial joints and to facilitate bone
 CC reconstruction. They can also be used to treat e.g. ulcers, lesions,
 CC injuries, burns, trauma, periodontal conditions, lacerations and other
 CC conditions. The invention is also useful in protein therapy. The present
 CC sequence is a linker peptide used in the invention.
 CX
 CQ Sequence 7 AA;
 CX
 Query Match 100.0%; Score 5; DB 24; Length 7;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Y 1 GGGGS 5
 b |
 1 GGGGS 5
 RESULT 98
 AE30034
 D AAE30034 standard; peptide; 7 AA.
 CX
 C AAE30034;
 CX
 T 24-FEB-2003 (first entry)
 CX
 E Peptide used to illustrate the method of the invention.
 CX
 W Entry localisation; light-generating fusion protein; LGP; diabetes;
 KW hypoxic tissue imaging; detection; cancer; hypoxia; heart disease;
 KW stroke.
 OS Unidentified.
 XX
 PN WO200275278-A2.
 XX
 PD 26-SEP-2002.
 XX
 PE 20-MAR-2002; 2002WO-US08864.
 XX
 PR 20-MAR-2001; 2001US-277425P.
 PR 20-MAR-2001; 2001US-277431P.
 PR 20-MAR-2001; 2001US-277440P.
 PR 09-NOV-2001; 2001US-332334P.
 PR 09-NOV-2001; 2001US-332493P.
 PR 09-NOV-2001; 2001US-345200P.
 PR 20-DEC-2001; 2001US-345298P.
 PR 20-DEC-2001; 2001US-345131P.
 PR 20-DEC-2001; 2001US-345132P.
 PR 19-MAR-2002; 2002US-0101662.
 XX
 PA (DAND) DANA FARBER CANCER INST INC.
 XX
 PI Kaelin WG, Livingston DM, Kim W;
 XX
 DR WPI; 2003-018815/01.
 XX
 PT Detecting localization of an entity e.g. hypoxic tissue, tumor, wound
 PT in a subject, by using a light-generating fusion protein having a
 PT ligand binding site and light-generating polypeptide moiety -
 XX
 PS Example 3; Fig 4A; 129pp; English.
 XX
 CC The invention relates to a method of detecting localisation of an
 CC entity in a subject. The method involves administering to the subject a
 CC light-generating fusion protein (LGP), or a cell expressing LGP, where
 CC LGP comprises a ligand binding site and a light-generating polypeptide
 CC moiety, and light generation of LGP changes upon binding of a ligand
 CC at the ligand binding site, allowing for co-localisation of LGP and an
 CC entity, and imaging localised LGP. The method is useful for detecting
 CC the localisation of an entity, such as a molecule, macromolecule,
 CC polymer, protein, antibody, protein complex, polysaccharide, nucleic
 CC acid, particle, inert material, organelle, cell, embryo, microorganism,
 CC bacteria, virus, fungus, prion, tumour tissue, cellular environment
 CC comprising damaged tissue, diseased tissue or hypoxic tissue, wound,
 CC organ, proliferating cell and pathogen in a subject. It is particularly
 CC useful for measuring prolly hydroxylase activity; for imaging hypoxic
 CC tissue and for detecting cancerous tissue. LGPs are useful for screening
 CC modulators of activity or latency of (or predisposition to) disorders
 CC such as hypoxia, cancer, diabetes, heart disease or stroke. The present
 CC sequence is a peptide used to illustrate the method of the invention.
 CC This sequence is used to illustrate the method of the invention.
 CX
 SQ Sequence 7 AA;
 CX
 Query Match 100.0%; Score 5; DB 24; Length 7;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGGS 5
 Db |
 1 GGGGS 5
 RESULT 99
 AAR86795
 ID AAR86795 standard; Peptide; 8 AA.
 XX
 AC AAR86795;
 XX
 DT 25-MAR-2003 (updated)
 DT 01-JUL-1996 (first entry)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 17:40:14 ; Search time 14.9017 Seconds
(without alignments)
495.460 Million cell updates/sec

Title: US-10-057-890A-31

Perfect score: 884

Sequence: 1 MKVSVAAALSCLMLVTALGSM.....GLNNCSSNRLDGHQRVHAA 157

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	379	42.9	352	1	CKR5_HUMAN
2	374	42.3	352	1	CKR5_PANTR
3	369	41.7	352	1	CKR5_PONPY
4	366	41.4	352	1	CKR5_PAPHA
5	365	41.3	352	1	CKR5_PYGB1
6	365	41.3	352	1	CKR5_PYGNE
7	365	41.3	352	1	CKR5_TRAFR
8	365	41.3	352	1	CKR5_TRAFR
9	364	41.2	352	1	CKR5_GORGO
10	363	41.1	352	1	CKR5_MACMU
11	361	40.8	352	1	CKR5_CERTO
12	359	40.6	352	1	CKR5_HYLLE
13	357	40.4	352	1	CKR5_CERAE
14	272	30.8	354	1	CKR5_MOUSE
15	249	28.2	354	1	CKR5_RAT
16	156.5	17.7	474	1	Z256_HUMAN
17	154	17.4	201	1	Z239_MOUSE
18	151.5	17.1	645	1	Z235_MOUSE
19	151.5	17.1	754	1	Z287_HUMAN
20	150.5	17.0	501	1	ZF96_MOUSE
21	146.5	16.6	458	1	Z239_HUMAN
22	144.5	16.3	759	1	Z287_MOUSE
23	144	16.3	698	1	Z234_HUMAN
24	140.5	15.9	604	1	Z305_HUMAN
25	140.5	15.9	744	1	Y362_HUMAN
26	140.5	15.9	803	1	Z226_HUMAN
27	140.5	15.9	913	1	Z228_HUMAN
28	140	15.8	578	1	Z192_HUMAN
29	139.5	15.8	614	1	ZF28_MOUSE
30	139.5	15.8	682	1	ZN45_HUMAN
31	139.5	15.8	751	1	Z184_HUMAN
32	139	15.7	810	1	Z33A_HUMAN
33	138.5	15.7	223	1	GL14_HUMAN

34 138.5 15.7 670 1 ZN16_HUMAN P17020 homo sapien
35 138.5 15.7 1029 1 Z197_HUMAN O14709 homo sapien
36 136.5 15.4 348 1 Z134_HUMAN P52741 homo sapien
37 136.5 15.4 659 1 Z304_HUMAN Q95CX3 homo sapien
38 135.5 15.3 643 1 ZN74_HUMAN O16587 homo sapien
39 133.5 15.1 636 1 ZF90_MOUSE O61967 mus musculus
40 132.5 15.0 642 1 ZN14_HUMAN P17017 homo sapien
41 131.5 14.9 367 1 Z211_HUMAN Q13398 homo sapien
42 131.5 14.9 446 1 ZN38_HUMAN P17036 homo sapien
43 131.5 14.9 506 1 Z157_HUMAN P51786 homo sapien
44 131.5 14.9 686 1 ZN07_HUMAN P17097 homo sapien
45 131 14.8 739 1 YD49_HUMAN Q9P2J8 homo sapien

ALIGNMENTS

RESULT 1
CKR5_HUMAN STANDARD; PRT; 352 AA.
AC P51681; O14692; O14693; O14695; O14696; O14697; O14698; O14699;
AC O14700; O14701; O14702; O14703; O14704; O14705; O14706; O14707;
AC O14708; O15538; Q9UPA4;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (CCR5)
DE (HIV-1 fusion co-receptor) (CHEMR13) (CD195 antigen).
GN CKR5 OR CMKBR5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96241590; PubMed=8639485;
RA Samson M., Labbe O., Mollereau C., Vassart G., Parmentier M.;
RT "Molecular cloning and functional expression of a new human
RT CC-chemokine receptor gene."
RL Biochemistry 35:3362-3367(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96291862; PubMed=8663314;
RA Rapport C.J., Gosling J., Schweichart V.L., Gray P.W., Charo I.F.;
RT "Molecular cloning and functional characterization of a novel human
RT CC chemokine receptor (CCR5) for RANTES, MIP-1beta, and MIP-1alpha."
RL J. Biol. Chem. 271:17161-17166(1996).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96295970; PubMed=8699119;
RA Combadiere C., Ahuja S.K., Tiffany H.L., Murphy P.M.;
RT "Cloning and functional expression of CC CKR5, a human monocyte CC
RT chemokine receptor selective for MIP-1(alpha), MIP-1(beta), and
RT RANTES."
RL J. Leukoc. Biol. 60:147-152(1996).
RN [4]
RP SEQUENCE FROM N.A.
RX McCombie W.R., Wilson R., Chen E., Gibbs R., Zuo L., Johnson D.,
RA Nhan M., Parnell L., Dedhia N., Ansari A., Mardis E., Schutz K.,
RA Gnoj L., la Bastide M., Kaplan N., Greco T., Touchman J.,
RA Muzny D., Chen C.N., Evans C., Fitzgerald M., See L.H., Tang M.,
RA Porcelsky K.A., Desilva U., Giacalone J., Pae A., Powell E.,
RA Solinsky K.A., Diaz-Perez S., Zhou X., Yu Y.,
RA Watanabe M., Doggett N., Garcia D., Sagripanti J.L.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=98001387; PubMed=9343222;
RA Kuhnmann S.E., Platt E.J., Kozak S.L., Kabat D.;
RT "Polymorphisms in the CCR5 genes of African green monkeys and mice
RT implicate specific amino acids in infections by simian and human
RT immunodeficiency viruses."
RL J. Virol. 71:8642-8656(1997).

[6] SEQUENCE FROM N.A., AND POLYMORPHISMS.
 MEDLINE=98022612; PubMed=9359654;
 Zhang L., Carruthers C.D., He T., Huang Y., Cao Y., Wang G., Hahn B.,
 Ho D.D.;
 "HIV type 1 subtypes, coreceptor usage, and CCR5 polymorphism.";
 AIDS Res. Hum. Retroviruses 13:1357-1366(1997).
 [7] SEQUENCE FROM N.A.
 MEDLINE=98049523; PubMed=9388201;
 Mummidi S., Ahuja S.S., McDaniel B.L., Ahuja S.K.;
 "The human CC chemokine receptor 5 (CCR5) gene. Multiple transcripts
 with 5'-end heterogeneity, dual promoter usage, and evidence for
 polymorphisms within the regulatory regions and noncoding exons.";
 J. Biol. Chem. 272:30662-30671(1997).
 [8] SEQUENCE FROM N.A., AND VARIANT ARG-178.
 Magierowska M., Barre-Sinoussi F., Issafras H., Theodorou I.,
 Debre P.;
 Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 [9] CHARACTERIZATION OF ITS HIV-1 RECEPTOR FUNCTION.
 MEDLINE=96260017; PubMed=8649511;
 Deng H., Liu R., Ellmeier W., Choe S., Unutmaz D., Burkhart M.,
 Di Marzio P., Broder C., Martin S.R., Hill C.M., Davis C.B.,
 Peiper S.C., Schall T.J., Littman D.R., Landau N.R.;
 "Identification of a major co-receptor for primary isolates of
 HIV-1.";
 Nature 381:661-666(1996).
 [10] CHARACTERIZATION OF ITS HIV-1 RECEPTOR FUNCTION.
 MEDLINE=96260018; PubMed=8649512;
 Dragic T., Litwin V., Allaway G.P., Martin S.R., Huang Y.,
 Nagashima K.A., Cayanan C., Maddon P.J., Koup R.A., Moore J.P.,
 Paxton W.A.;
 "HIV-1 entry into CD4+ cells is mediated by the chemokine receptor.
 CC-CR5.";
 Nature 381:667-673(1996).
 [11] SULFATION.
 MEDLINE=99189752; PubMed=10089882;
 Farzan M., Mirzabekov T., Kolchinsky P., Wyatt R., Cayabyab M.,
 Gerard N.P., Gerard C., Sodroski J., Choe H.;
 "Tyrosine sulfation of the amino terminus of CCR5 facilitates HIV-1
 entry.";
 Cell 96:667-676(1999).
 C-1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 DIFFERENTIATION. ACTS AS CO-RECEPTOR WITH CD4 FOR PRIMARY NON-
 SYNCYTUM-INDUCING STRAINS (NSI) (MACROPHAGE-TROPIC) OF HIV-1
 VIRUS. IT PROMOTES ENV-MEDIATED FUSION OF THE VIRUS.
 C-2- SUBCELLULAR LOCATION: Integral membrane protein.
 C-3- TISSUE SPECIFICITY: FOUND IN PROMYELOCYTIC CELLS.
 C-4- PTM: SULFATION CONTRIBUTES TO THE EFFICIENCY OF HIV-1 ENTRY.
 C-5- PTM: MODIFIED BY O-LINKED GLYCOSYLATION, BUT NOT BY N-LINKED
 GLYCOSYLATION.
 C-6- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 C-7- This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 C-8- EMBL: X31482; CAA62796.1; -
 C-9- EMBL: U54994; AAC50598.1; -
 C-10- EMBL: U57840; AAB10701.1; -
 C-11- EMBL: U95626; AAB57793.1; -
 C-12- EMBL: U83326; AAC51797.1; -

DR EMBL; AF011500; AAB65700.1; -
 DR EMBL; AF011501; AAB65701.1; -
 DR EMBL; AF011502; AAB65702.1; -
 DR EMBL; AF011503; AAB65703.1; -
 DR EMBL; AF011505; AAB65705.1; -
 DR EMBL; AF011506; AAB65706.1; -
 DR EMBL; AF011507; AAB65707.1; -
 DR EMBL; AF011508; AAB65708.1; -
 DR EMBL; AF011509; AAB65709.1; -
 DR EMBL; AF011510; AAB65710.1; -
 DR EMBL; AF011511; AAB65711.1; -
 DR EMBL; AF011512; AAB65712.1; -
 DR EMBL; AF011513; AAB65713.1; -
 DR EMBL; AF011514; AAB65714.1; -
 DR EMBL; AF011515; AAB65715.1; -
 DR EMBL; AF011516; AAB65716.1; -
 DR EMBL; AF011517; AAB65717.1; -
 DR EMBL; AF011518; AAB65718.1; -
 DR EMBL; AF011519; AAB65719.1; -
 DR EMBL; AF011520; AAB65720.1; -
 DR EMBL; AF011521; AAB65721.1; -
 DR EMBL; AF011522; AAB65722.1; -
 DR EMBL; AF011523; AAB65723.1; -
 DR EMBL; AF011524; AAB65724.1; -
 DR EMBL; AF011525; AAB65725.1; -
 DR EMBL; AF011526; AAB65726.1; -
 DR EMBL; AF011527; AAB65727.1; -
 DR EMBL; AF011528; AAB65728.1; -
 DR EMBL; AF011529; AAB65729.1; -
 DR EMBL; AF011530; AAB65730.1; -
 DR EMBL; AF011531; AAB65731.1; -
 DR EMBL; AF011532; AAB65732.1; -
 DR EMBL; AF011533; AAB65733.1; -
 DR EMBL; AF011534; AAB65734.1; -
 DR EMBL; AF011535; AAB65735.1; -
 DR EMBL; AF011536; AAB65736.1; -
 DR EMBL; AF011537; AAB65737.1; -
 DR EMBL; AF011538; AAB65738.1; -
 DR EMBL; AF011539; AAB65739.1; -
 DR EMBL; AF011540; AAB65740.1; -
 DR EMBL; AF011541; AAB65741.1; -
 DR EMBL; AF011542; AAB65742.1; -
 DR EMBL; AF011543; AAB65743.1; -
 DR EMBL; AF011544; AAB65744.1; -
 DR EMBL; AF011545; AAB65745.1; -
 DR EMBL; AF011546; AAB65746.1; -
 DR EMBL; AF011547; AAB65747.1; -
 DR EMBL; AF011548; AAB65748.1; -
 DR EMBL; AF011549; AAB65749.1; -
 DR EMBL; AF011550; AAB65750.1; -
 DR EMBL; AF011551; AAB65751.1; -
 DR EMBL; AF011552; AAB65752.1; -
 DR EMBL; AF011553; AAB65753.1; -
 DR EMBL; AF011554; AAB65754.1; -
 DR EMBL; AF011555; AAB65755.1; -
 DR EMBL; AF011556; AAB65756.1; -
 DR EMBL; AF011557; AAB65757.1; -
 DR EMBL; AF011558; AAB65758.1; -
 DR EMBL; AF011559; AAB65759.1; -
 DR EMBL; AF011560; AAB65760.1; -
 DR EMBL; AF011561; AAB65761.1; -
 DR EMBL; AF011562; AAB65762.1; -
 DR EMBL; AF011563; AAB65763.1; -
 DR EMBL; AF011564; AAB65764.1; -
 DR EMBL; AF011565; AAB65765.1; -
 DR EMBL; AF011566; AAB65766.1; -
 DR EMBL; AF011567; AAB65767.1; -
 DR EMBL; AF011568; AAB65768.1; -
 DR EMBL; AF011569; AAB65769.1; -
 DR EMBL; AF011570; AAB65770.1; -
 DR EMBL; AF011571; AAB65771.1; -
 DR EMBL; AF011572; AAB65772.1; -
 DR EMBL; AF011573; AAB65773.1; -
 DR EMBL; AF011574; AAB65774.1; -
 DR EMBL; AF011575; AAB65775.1; -
 DR EMBL; AF011576; AAB65776.1; -
 DR EMBL; AF011577; AAB65777.1; -
 DR EMBL; AF011578; AAB65778.1; -
 DR EMBL; AF011579; AAB65779.1; -
 DR EMBL; AF011580; AAB65780.1; -
 DR EMBL; AF011581; AAB65781.1; -
 DR EMBL; AF011582; AAB65782.1; -
 DR EMBL; AF011583; AAB65783.1; -
 DR EMBL; AF011584; AAB65784.1; -
 DR EMBL; AF011585; AAB65785.1; -
 DR EMBL; AF011586; AAB65786.1; -
 DR EMBL; AF011587; AAB65787.1; -
 DR EMBL; AF011588; AAB65788.1; -
 DR EMBL; AF011589; AAB65789.1; -
 DR EMBL; AF011590; AAB65790.1; -
 DR EMBL; AF011591; AAB65791.1; -
 DR EMBL; AF011592; AAB65792.1; -
 DR EMBL; AF011593; AAB65793.1; -
 DR EMBL; AF011594; AAB65794.1; -
 DR EMBL; AF011595; AAB65795.1; -
 DR EMBL; AF011596; AAB65796.1; -
 DR EMBL; AF011597; AAB65797.1; -
 DR EMBL; AF011598; AAB65798.1; -
 DR EMBL; AF011599; AAB65799.1; -
 DR EMBL; AF011600; AAB65800.1; -
 DR EMBL; AF011601; AAB65801.1; -
 DR EMBL; AF011602; AAB65802.1; -
 DR EMBL; AF011603; AAB65803.1; -
 DR EMBL; AF011604; AAB65804.1; -
 DR EMBL; AF011605; AAB65805.1; -
 DR EMBL; AF011606; AAB65806.1; -
 DR EMBL; AF011607; AAB65807.1; -
 DR EMBL; AF011608; AAB65808.1; -
 DR EMBL; AF011609; AAB65809.1; -
 DR EMBL; AF011610; AAB65810.1; -
 DR EMBL; AF011611; AAB65811.1; -
 DR EMBL; AF011612; AAB65812.1; -
 DR EMBL; AF011613; AAB65813.1; -
 DR EMBL; AF011614; AAB65814.1; -
 DR EMBL; AF011615; AAB65815.1; -
 DR EMBL; AF011616; AAB65816.1; -
 DR EMBL; AF011617; AAB65817.1; -
 DR EMBL; AF011618; AAB65818.1; -
 DR EMBL; AF011619; AAB65819.1; -
 DR EMBL; AF011620; AAB65820.1; -
 DR EMBL; AF011621; AAB65821.1; -
 DR EMBL; AF011622; AAB65822.1; -
 DR EMBL; AF011623; AAB65823.1; -
 DR EMBL; AF011624; AAB65824.1; -
 DR EMBL; AF011625; AAB65825.1; -
 DR EMBL; AF011626; AAB65826.1; -
 DR EMBL; AF011627; AAB65827.1; -
 DR EMBL; AF011628; AAB65828.1; -
 DR EMBL; AF011629; AAB65829.1; -
 DR EMBL; AF011630; AAB65830.1; -
 DR EMBL; AF011631; AAB65831.1; -
 DR EMBL; AF011632; AAB65832.1; -
 DR EMBL; AF011633; AAB65833.1; -
 DR EMBL; AF011634; AAB65834.1; -
 DR EMBL; AF011635; AAB65835.1; -
 DR EMBL; AF011636; AAB65836.1; -
 DR EMBL; AF011637; AAB65837.1; -
 DR EMBL; AF011638; AAB65838.1; -
 DR EMBL; AF011639; AAB65839.1; -
 DR EMBL; AF011640; AAB65840.1; -
 DR EMBL; AF011641; AAB65841.1; -
 DR EMBL; AF011642; AAB65842.1; -
 DR EMBL; AF011643; AAB65843.1; -
 DR EMBL; AF011644; AAB65844.1; -
 DR EMBL; AF011645; AAB65845.1; -
 DR EMBL; AF011646; AAB65846.1; -
 DR EMBL; AF011647; AAB65847.1; -
 DR EMBL; AF011648; AAB65848.1; -
 DR EMBL; AF011649; AAB65849.1; -
 DR EMBL; AF011650; AAB65850.1; -
 DR EMBL; AF011651; AAB65851.1; -
 DR EMBL; AF011652; AAB65852.1; -
 DR EMBL; AF011653; AAB65853.1; -
 DR EMBL; AF011654; AAB65854.1; -
 DR EMBL; AF011655; AAB65855.1; -
 DR EMBL; AF011656; AAB65856.1; -
 DR EMBL; AF011657; AAB65857.1; -
 DR EMBL; AF011658; AAB65858.1; -
 DR EMBL; AF011659; AAB65859.1; -
 DR EMBL; AF011660; AAB65860.1; -
 DR EMBL; AF011661; AAB65861.1; -
 DR EMBL; AF011662; AAB65862.1; -
 DR EMBL; AF011663; AAB65863.1; -
 DR EMBL; AF011664; AAB65864.1; -
 DR EMBL; AF011665; AAB65865.1; -
 DR EMBL; AF011666; AAB65866.1; -
 DR EMBL; AF011667; AAB65867.1; -
 DR EMBL; AF011668; AAB65868.1; -
 DR EMBL; AF011669; AAB65869.1; -
 DR EMBL; AF011670; AAB65870.1; -
 DR EMBL; AF011671; AAB65871.1; -
 DR EMBL; AF011672; AAB65872.1; -
 DR EMBL; AF011673; AAB65873.1; -
 DR EMBL; AF011674; AAB65874.1; -
 DR EMBL; AF011675; AAB65875.1; -
 DR EMBL; AF011676; AAB65876.1; -
 DR EMBL; AF011677; AAB65877.1; -
 DR EMBL; AF011678; AAB65878.1; -
 DR EMBL; AF011679; AAB65879.1; -
 DR EMBL; AF011680; AAB65880.1; -
 DR EMBL; AF011681; AAB65881.1; -
 DR EMBL; AF011682; AAB65882.1; -
 DR EMBL; AF011683; AAB65883.1; -
 DR EMBL; AF011684; AAB65884.1; -
 DR EMBL; AF011685; AAB65885.1; -
 DR EMBL; AF011686; AAB65886.1; -
 DR EMBL; AF011687; AAB65887.1; -
 DR EMBL; AF011688; AAB65888.1; -
 DR EMBL; AF011689; AAB65889.1; -
 DR EMBL; AF011690; AAB65890.1; -
 DR EMBL; AF011691; AAB65891.1; -
 DR EMBL; AF011692; AAB65892.1; -
 DR EMBL; AF011693; AAB65893.1; -
 DR EMBL; AF011694; AAB65894.1; -
 DR EMBL; AF011695; AAB65895.1; -
 DR EMBL; AF011696; AAB65896.1; -
 DR EMBL; AF011697; AAB65897.1; -
 DR EMBL; AF011698; AAB65898.1; -
 DR EMBL; AF011699; AAB65899.1; -
 DR EMBL; AF011700; AAB65900.1; -
 DR EMBL; AF011701; AAB65901.1; -
 DR EMBL; AF011702; AAB65902.1; -
 DR EMBL; AF011703; AAB65903.1; -
 DR EMBL; AF011704; AAB65904.1; -
 DR EMBL; AF011705; AAB65905.1; -
 DR EMBL; AF011706; AAB65906.1; -
 DR EMBL; AF011707; AAB65907.1; -
 DR EMBL; AF011708; AAB65908.1; -
 DR EMBL; AF011709; AAB65909.1; -
 DR EMBL; AF011710; AAB65910.1; -
 DR EMBL; AF011711; AAB65911.1; -
 DR EMBL; AF011712; AAB65912.1; -
 DR EMBL; AF011713; AAB65913.1; -
 DR EMBL; AF011714; AAB65914.1; -
 DR EMBL; AF011715; AAB65915.1; -
 DR EMBL; AF011716; AAB65916.1; -
 DR EMBL; AF011717; AAB65917.1; -
 DR EMBL; AF011718; AAB65918.1; -
 DR EMBL; AF011719; AAB65919.1; -
 DR EMBL; AF011720; AAB65920.1; -
 DR EMBL; AF011721; AAB65921.1; -
 DR EMBL; AF011722; AAB65922.1; -
 DR EMBL; AF011723; AAB65923.1; -
 DR EMBL; AF011724; AAB65924.1; -
 DR EMBL; AF011725; AAB65925.1; -
 DR EMBL; AF011726; AAB65926.1; -
 DR EMBL; AF011727; AAB65927.1; -
 DR EMBL; AF011728; AAB65928.1; -
 DR EMBL; AF011729; AAB65929.1; -
 DR EMBL; AF011730; AAB65930.1; -
 DR EMBL; AF011731; AAB65931.1; -
 DR EMBL; AF011732; AAB65932.1; -
 DR EMBL; AF011733; AAB65933.1; -
 DR EMBL; AF011734; AAB65934.1; -
 DR EMBL; AF011735; AAB65935.1; -
 DR EMBL; AF011736; AAB65936.1; -
 DR EMBL; AF011737; AAB65937.1; -
 DR EMBL; AF011738; AAB65938.1; -
 DR EMBL; AF011739; AAB65939.1; -
 DR EMBL; AF011740; AAB65940.1; -
 DR EMBL; AF011741; AAB65941.1; -
 DR EMBL; AF011742; AAB65942.1; -
 DR EMBL; AF011743; AAB65943.1; -
 DR EMBL; AF011744; AAB65944.1; -
 DR EMBL; AF011745; AAB65945.1; -
 DR EMBL; AF011746; AAB65946.1; -
 DR EMBL; AF011747; AAB65947.1; -
 DR EMBL; AF011748; AAB65948.1; -
 DR EMBL; AF011749; AAB65949.1; -
 DR EMBL; AF011750; AAB65950.1; -
 DR EMBL; AF011751; AAB65951.1; -
 DR EMBL; AF011752; AAB65952.1; -
 DR EMBL; AF011753; AAB65953.1; -
 DR EMBL; AF011754; AAB65954.1; -
 DR EMBL; AF011755; AAB65955.1; -
 DR EMBL; AF011756; AAB65956.1; -
 DR EMBL; AF011757; AAB65957.1; -
 DR EMBL; AF011758; AAB65958.1; -
 DR EMBL; AF011759; AAB65959.1; -
 DR EMBL; AF011760; AAB65960.1; -
 DR EMBL; AF011761; AAB65961.1; -
 DR EMBL; AF011762; AAB65962.1; -
 DR EMBL; AF011763; AAB65963.1; -
 DR EMBL; AF011764; AAB65964.1; -
 DR EMBL; AF011765; AAB65965.1; -
 DR EMBL; AF011766; AAB65966.1; -
 DR EMBL; AF011767; AAB65967.1; -
 DR EMBL; AF011768; AAB65968.1; -
 DR EMBL; AF011769; AAB65969.1; -
 DR EMBL; AF011770; AAB65970.1; -
 DR EMBL; AF011771; AAB65971.1; -
 DR EMBL; AF011772; AAB65972.1; -
 DR EMBL; AF011773; AAB65973.1; -
 DR EMBL; AF011774; AAB65974.1; -
 DR EMBL; AF011775; AAB65975.1; -
 DR EMBL; AF011776; AAB65976.1; -
 DR EMBL; AF011777; AAB65977.1; -
 DR EMBL; AF011778; AAB65978.1; -
 DR EMBL; AF011779; AAB65979.1; -
 DR EMBL; AF011780; AAB65980.1; -
 DR EMBL; AF011781; AAB65981.1; -
 DR EMBL; AF011782; AAB65982.1; -
 DR EMBL; AF011783; AAB65983.1; -
 DR EMBL; AF011784; AAB65984.1; -
 DR EMBL; AF011785; AAB65985.1; -
 DR EMBL; AF011786; AAB65986.1; -
 DR EMBL; AF011787; AAB65987.1; -
 DR EMBL; AF011788; AAB65988.1; -
 DR EMBL; AF011789; AAB65989.1; -
 DR EMBL; AF011790; AAB65990.1; -
 DR EMBL; AF011791; AAB65991.1; -
 DR EMBL; AF011792; AAB65992.1; -
 DR EMBL; AF011793; AAB65993.1; -
 DR EMBL; AF011794; AAB65994.1; -
 DR EMBL; AF011795; AAB65995.1; -
 DR EMBL; AF011796; AAB65996.1; -
 DR EMBL; AF011797; AAB65997.1; -
 DR EMBL; AF011798; AAB65998.1; -
 DR EMBL; AF011799; AAB65999.1; -
 DR EMBL; AF011800; AAB66000.1; -

Query Match 42.9%; Score 379; DB 1; Length 352;
 Best Local Similarity 35.0%; Pred. No. 1.3e-29;
 Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

QY 20 MDYQSSPIYDINYYTSPCKINVKQIAA----- 49
 DB 1 MDYQSSPIYDINYYTSPCKINVKQIAARLLPLSLVIFGFGVGNMLVILINCKR 60
 QY 50 -----YKGLC-----AAAQDFGNTMCOHORVGHGHHHSYKCG--- 83
 DB 61 LKSMTDIVLLNLAISSDLFFLLTPFWAHYAAQWDFGNTMC--QLLTGLYFIFGFGSIIFF 118
 QY 84 -----LCTRSQKGLHYTC 97
 DB 119 ILLTIDRYLAIVHVAFFALKARTVTVGVTSVITWWAVFASLPGLIIFTRSQKGLHYTC 178
 QY 98 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 124
 DB 179 SSHFPYSQYQFWKNFQTLKIIVGLVPLVWVICYSGLKTLRCRNEKXKRAVRLIF 238
 QY 125 -----GGSYKGLC-----QEPFGLNCCSSNRDLGHORV 154
 DB 239 TIMIVYFLFMAPYINIVLLNTFQEPFGLNCCSSNRDLQAMQV 281

RESULT 2
 CCR5_PANTR
 ID CCR5_PANTR STANDARD; PRT; 352 AA.
 AC P56440; O02778;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CCR5).
 GN CCR5 OR CMKR5.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 OX NCBI TaxID:9598;
 RN [1] _
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97268687; PubMed=9108095;
 RA Edinger A.B., Amodee A., Miller K., Doranz B.J., Endres M.,
 RA Sharon M., Samson M., Lu Z.H., Clements J.E., Murphy-Corb M.,
 RA Peiper S.C., Parmentier M., Broder C.C., Doms R.W.;
 RT "Differential utilization of CCR5 by macrophage and T cell tropic
 RT simian immunodeficiency virus strains";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:4005-4010(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Zimmerman P.A., Buckler-White A., Alkhatib G.;
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98022612; PubMed=9359654;
 RA Zhang L., Carruthers C.D., He T., Huang Y., Cao Y., Wang G., Hahn B.,
 RA Ho D.D.;
 RT "HIV type 1 subtypes, coreceptor usage, and CCR5 polymorphism";
 RL AIDS Res. Hum. Retroviruses 13:1357-1366(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97426118; PubMed=9282822;
 RA Zacharova V., Zachar V., Goustin A.S.;
 RT "Sequence of chemokine receptor gene CCR5 in chimpanzees, a natural
 RT HIV type 1 host";
 RL AIDS Res. Hum. Retroviruses 13:1159-1161(1997).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98090115; PubMed=9430250;
 RA Pretet J.-L., Zerbib A., Girard M., Guillet J.-G., Butor C.;
 RT "Chimpanzee CXCR4 and CCR5 act as coreceptors for HIV type 1";
 RL AIDS Res. Hum. Retroviruses 13:1583-1587(1997).
 CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,

CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 CC DIFFERENTIATION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF005663; AAB62557.1; -
 CC EMBL; U94329; AAB58446.1; -
 CC EMBL; AF011542; AAB65742.1; -
 CC EMBL; U97666; AAC51670.1; -
 CC EMBL; AF011540; AAB65740.1; -
 CC EMBL; U89797; AAC03171.1; -
 CC InterPro; IPR000276; GPCR_Rhodopsn.
 CC Pfam; PF00001; 7tm_1; 1.
 CC PRINTS; PR00237; GPCRHOPOPSN.
 CC PROSITE; PS00237; G-PROTEIN RECP Fl_1; 1.
 CC PROSITE; PS0262; G-PROTEIN RECP Fl_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
 FT DOMAIN 1 30
 FT TRANSMEM 31 58
 FT DOMAIN 59 68
 FT TRANSMEM 69 89
 FT DOMAIN 90 102
 FT TRANSMEM 103 124
 FT DOMAIN 125 141
 FT TRANSMEM 142 166
 FT DOMAIN 167 198
 FT TRANSMEM 199 218
 FT DOMAIN 219 235
 FT TRANSMEM 236 260
 FT DOMAIN 261 277
 FT TRANSMEM 278 301
 FT DOMAIN 302 352
 FT DISULFID 101 178
 FT MOD_RES 3 3
 FT MOD_RES 10 10
 FT MOD_RES 14 14
 FT MOD_RES 15 15
 FT CAREQHD 268 268
 FT CONFLICT 123 123
 SQ SEQUENCE 352 AA; 40539 MW; 4A33B69B80FE34C CRC64;
 Query Match 42.3%; Score 374; DB 1; Length 352;
 Best Local Similarity 34.6%; Pred. No. 3.9e-29;
 Matches 98; Conservative 8; Mismatches 27; Indels 150; Gaps 7;

QY 20 MDYQSSPIYDINYYTSPCKINVKQIAA----- 49
 DB 1 MDYQSSPIYDINYYTSPCKINVKQIAARLLPLSLVIFGFGVGNMLVILINCKR 60
 QY 50 -----YKGLC-----AAAQDFGNTMCOHORVGHGHHHSYKCG--- 83
 DB 61 LKSMTDIVLLNLAISSDLFFLLTPFWAHYAAQWDFGNTMC--QLLTGLYFIFGFGSIIFF 118
 QY 84 -----LCTRSQKGLHYTC 97
 DB 119 ILLTIDRYLAIVHVAFFALKARTVTVGVTSVITWWAVFASLPGLIIFTRSQKGLHYTC 178
 QY 98 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 124
 DB 179 SSHFPYSQYQFWKNFQTLKIIVGLVPLVWVICYSGLKTLRCRNEKXKRAVRLIF 238
 QY 125 -----GGSYKGLC-----QEPFGLNCCSSNRDLGHORV 154

239 TIMIVYFLWAPYNTIVLLNTFQEFGLNCCSSNRLDQAMQV 281

Db 1 MDYQVSSPTDYDIDYVTSPECQKINVKQIAARLLPPLSLVFIQVGNMLVILILINCKR 60
 QY 50 -----YKGLC-----AAQWDFGNTCMQHORVGHGHHHSYKCG--- 83
 Db 61 LKSMTDIVLLANLAISDLFFLLTVPFWAHYAAQWDFGNTCM---QLLTGLYFIQFSGIFF 118
 QY 84 -----LCTRSQKGLHYTC 97
 Db 119 IILLTIDRYLAIVHAFVAKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKGLHYTC 178
 QY 98 SHGPPYSQYQVWKNFQTLKI-----HORVHG----- 124
 Db 179 SHGPPYSQYQVWKNFQTLKIIVILGLVPLLVVVICSYGILATLLRCRNEKKRRAVRLIF 238
 QY 125 -----GGSYKGLC-----QEFGLNCCSSNRLDGHORV 154
 Db 239 TIMIVYFLWAPYNTIVLLNTFQEFGLNCCSSNRLDQAMQV 281

RESULT 4

CKRS_PAPHA
 ID CKRS_PAPHA STANDARD; PRT; 352 AA.
 AC P56441;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DB C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (CCR5).
 GN CCR5 OR CKMR5.
 OS Papio hamadryas (Hamadryas baboon), and
 OS Papio anubis (Olive baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
 OC Cercopitheciinae; Papio.
 OX NCBI_TaxID:9557, 9555;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=P.hamadryas;
 RX MEDLINE=97268687; PubMed=9108095;
 RA Edinger A.L., Amedee A., Miller K., Doranz B.J., Endres M.,
 RA Sharron M., Samson M., Lu Z.H., Clements J.E., Murphy-Corb M.,
 RA Peiper S.C., Parmentier M., Broder C.C., Doms R.W.;
 RT "Differential utilization of CCR5 by macrophage and T cell tropic
 RT simian immunodeficiency virus strains.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:4005-4010(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=P.hamadryas;
 RX MEDLINE=99210133; PubMed=10195758;
 RA Saksena N.K., Wang B., Novembre F.J., Bolton W., Smit T.K., Lal R.B.;
 RT "Species-specific changes in the CCR5 gene from African and Asian
 RT nonhuman primates.";
 RL AIDS Res. Hum. Retroviruses 15:479-483(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=P.anubis;
 RX Benton P.A., Timanus D.K., Shearer M.H., Lee D.R., Kennedy R.C.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 CC DIFFERENTIATION.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)

Db 239 TIMIVYFLWAPYNTIVLLNTFQEFGLNCCSSNRLDQAMQV 281
 ID CKRS_PONPY STANDARD; PRT; 352 AA.
 AC O97881;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DB C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (CCR5).
 GN Pongo pygmaeus (Orangutan).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pongo.
 OX NCBI_TaxID=9600;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=99416438; PubMed=10486970;
 RA Zhang Y.-W., Ryder O.A., Zhang Y.-P.;
 RL "Sequence evolution of the CCR5 chemokine receptor gene in primates.";
 RL Mol. Biol. Evol. 16:1145-1154(1999).
 CC -!- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 CC DIFFERENTIATION.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)

EMBL: AF075446; RAD19858.1; -;
 InterPro: IPR000276; GPCR_Rhodopsin.
 Pfam: PF00001; 7tm1; 1.
 PRINTS: PR00237; GPCRHHODOPSN.
 PROSITE: PS00237; G-PROTEIN RECEPTOR FL1; 1.
 PROSITE: PS00262; G-PROTEIN RECEPTOR FL2; 1.
 G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
 DOMAIN 1 30
 TRANSMEM 31 58
 DOMAIN 59 68
 TRANSMEM 69 89
 DOMAIN 90 102
 TRANSMEM 103 124
 DOMAIN 125 141
 TRANSMEM 142 166
 DOMAIN 167 198
 TRANSMEM 199 218
 DOMAIN 219 235
 TRANSMEM 236 260
 DOMAIN 261 277
 TRANSMEM 278 301
 DOMAIN 302 352
 DISULFID 101 178
 MOD_RES 3 3
 MOD_RES 10 10
 MOD_RES 14 14
 MOD_RES 15 15
 SEQUENCE 352 AA; 40527 MW; P4E2P47135AF658A CRC64;
 Query Match 41.7%; Score 369; DB 1; Length 352;
 Best Local Similarity 34.3%; Pred. No. 1.2e-28;
 Matches 97; Conservative 8; Mismatches 28; Indels 150; Gaps 7;
 20 MDYQVSSPTDYDIDYVTSPECQKINVKQIAARLLPPLSLVFIQVGNMLVILILINCKR 49

```

DR EMBL: AF005658; AAB62552.1; -
DR EMBL: AF105287; AAD20556.1; -
DR EMBL: AF105288; AAD20557.1; -
DR EMBL: AF105289; AAD20558.1; -
DR EMBL: AF105290; AAD20559.1; -
DR EMBL: AF023452; AAC63830.1; -
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm.1; 1.
DR PRINTS: PR00237; GPCRHOOPS.
DR PROSITE: PS00237; G-PROTEIN RECF_F1.1; 1.
DR PROSITE: PS0262; G-PROTEIN RECF_F1.2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; sulfation.
FT DOMAIN 1 30
FT TRANSMEM 31 58
FT DOMAIN 59 68
FT TRANSMEM 69 89
FT DOMAIN 90 102
FT TRANSMEM 103 124
FT DOMAIN 125 141
FT TRANSMEM 142 166
FT DOMAIN 167 198
FT TRANSMEM 199 218
FT DOMAIN 219 235
FT TRANSMEM 236 260
FT DOMAIN 261 277
FT TRANSMEM 278 301
FT DOMAIN 302 352
FT DISULFID 101 178
FT MOD_RES 3 3
FT MOD_RES 10 10
FT MOD_RES 14 14
FT MOD_RES 15 15
FT CARBOHYD 268 268
FT SEQUENCE 352 AA; 40489 MW; 5E1504A9BA1FE8B2 CRC64;
SQ
Query Match 41.4%; Score 366; DB 1; Length 352;
Best Local Similarity 33.9%; Pred. No. 2.3e-28;
Matches 96; Conservative 9; Mismatches 28; Indels 150; Gaps 7;

QY 20 MDYQSSPYDINYITSEPCQKINVKQIAA----- 49
DB 1 MDYQSSPYDIDYITSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNILVLLINCKR 60
QY 50 -----YKGLC-----AAQWDFGNTMCOHORVGHGHSYKCG--- 83
DB 61 LKSMTDIYLLNLAISDLFLLPVFPMAHYAAQWDFGNTMC--QLTLGYTFGFGSIF 118
QY 84 -----YKGLC-----LCTRSQKGLHYTC 97
DB 119 IILLTIDRYLAIVHAVFALKARTVTGVVTSVITWVAVFASLPGLIIFTRSQREGLHYTC 178
QY 98 SSHFPYSQYQFWKQFOTLKI-----HORVHG----- 124
DB 179 SSHFPYSQYQFWKQFOTLKIIVGLVPLVMVICYSGLKTLRCRNEKKHRAVRLIF 238
QY 125 -----GGSYKGLC-----QEFFGLNCCSSNRLDQMOV 281
DB 239 TIMIVFLFPADYINVLINTQEFFGLNCCSSNRLDQMOV 281

RESULT 5
ID CKR5_PYGBI STANDARD; PRT; 352 AA.
AC O97880;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CCR5).
GN CCR5 OR CMKBR5.
OS Pygathrix bieti (Black snub-nosed monkey) (Rhinopithecus bieti).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Butheria; Primates; Catarrhini; Cercopithecoidea; Colobinae;
OC Pygathrix.

```

```

OX NCBI_TaxID=61621;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99416438; PubMed=10486970;
RT Zhang Y.-W., Ryder O.A., Zhang Y.-P.;
RL Mol. Biol. Evol. 16:1145-1154 (1999)
CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
CC DIFFERENTIATION.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF075445; AAD19857.1; -
CC InterPro: IPR000276; GPCR_Rhodopsn.
CC Pfam: PF00001; 7tm.1; 1.
CC PRINTS: PR00237; GPCRHOOPS.
CC PROSITE: PS00237; G-PROTEIN RECF_F1.1; 1.
CC PROSITE: PS0262; G-PROTEIN RECF_F1.2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
FT DOMAIN 1 30
FT TRANSMEM 31 58
FT DOMAIN 59 68
FT TRANSMEM 69 89
FT DOMAIN 90 102
FT TRANSMEM 103 124
FT DOMAIN 125 141
FT TRANSMEM 142 166
FT DOMAIN 167 198
FT TRANSMEM 199 218
FT DOMAIN 219 235
FT TRANSMEM 236 260
FT DOMAIN 261 277
FT TRANSMEM 278 301
FT DOMAIN 302 352
FT DISULFID 101 178
FT MOD_RES 3 3
FT MOD_RES 10 10
FT MOD_RES 14 14
FT MOD_RES 15 15
FT SEQUENCE 352 AA; 40585 MW; 4366F142730F938F CRC64;
SQ
Query Match 41.3%; Score 365; DB 1; Length 352;
Best Local Similarity 33.6%; Pred. No. 2.9e-28;
Matches 95; Conservative 10; Mismatches 28; Indels 150; Gaps 7;

QY 20 MDYQSSPYDINYITSEPCQKINVKQIAA----- 49
DB 1 MDYQSSPYDIDYITSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNILVLLINCKR 60
QY 50 -----YKGLC-----AAQWDFGNTMCOHORVGHGHSYKCG--- 83
DB 61 LKSMTDIYLLNLAISDLFLLPVFPMAHYAAQWDFGNTMC--QLTLGYTFGFGSIF 118
QY 84 -----YKGLC-----LCTRSQKGLHYTC 97
DB 119 IILLTIDRYLAIVHAVFALKARTVTGVVTSVITWVAVFASLPGLIIFTRSQREGLHYTC 178
QY 98 SSHFPYSQYQFWKQFOTLKI-----HORVHG----- 124
DB 179 SSHFPYSQYQFWKQFOTLKIIVGLVPLVMVICYSGLKTLRCRNEKKHRAVRLIF 238
QY 125 -----GGSYKGLC-----QEFFGLNCCSSNRLDGHQVR 154

```

Query Match 41.3%; Score 365; DB 1; Length 352;
Best Local Similarity 33.6%; Pred. No. 2.9e-28;
Matches 95; Conservative 10; Mismatches 28; Indels 150; Gaps 7;

FT TRANSMEM 199 218 5 (POTENTIAL).
 FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 236 260 6 (POTENTIAL).
 FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 278 301 7 (POTENTIAL).
 FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 101 178 BY SIMILARITY.
 FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
 FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
 FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
 FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
 SQ SEQUENCE 352 AA; 40496 MW; 4366F148C255938F CRC64;

Query Match 41.3%; Score 365; DB 1; Length 352;
 Best Local Similarity 33.6%; Pred. No. 2.9e-28;
 Matches 95; Conservative 10; Mismatches 28; Indels 150; Gaps 7;

QY 20 MDYQVSSPIYDINYTSEPCQKINVKQIAA----- 49
 Db 1 MDYQVSSPTDYDYTSEPCQKINVKQIAARLLPLYSLVFIFGVGNILVVLINCKR 60
 QY 50 -----YKGLC-----AAQWDFGNTMCOHQRVHGHHSYKCG--- 83
 Db 61 LKSMTDIYLLNLAIISDLFFLLTPVFWAHYAAQWDFGNTMC--QLLTGLYPIGFSGIFF 118
 QY 84 ----- 124
 Db 119 IILLTDIDRYLAIVHAFALKARTVTGVVTSVITWVAVFASLPGIIFTRSQREGHLHYTC 178
 QY 98 SSHFPYSQYQFKNFOTLKI-----LCTRSQKEGLHYTC 97
 Db 179 SSHFPYSQYQFKNFOTLKIIVGLVPLLVNVICYSGLKTLRCRNEKKHRAVRLIF 238
 QY 125 -----GGSYKCGLC-----QEFFGLNCSNRLDGHQRV 154
 Db 239 TIMIVYFLFWAPYNIIVLLNTFOEFFGLNCSNRLDQAMQV 281

RESULT 8

CKR5 TRAPH
 ID CKR5 TRAPH STANDARD; PRT; 352 AA.
 AC O97879;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CCR5).
 GN CKR5 OR CMKRS.
 OS Trachypithecus phayrei (Phayre's leaf monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Colobinae;
 OC Trachypithecus.
 OX NCBI_TaxID=61618;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99416438; PubMed=10486970;
 RA Zhang Y.-W., Ryder O.A., Zhang Y.-P.;
 RL Mol. Biol. Evol. 16:1145-1154(1999).
 CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 CC DIFFERENTIATION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AF075443; AAD19855.1; -.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PRINTS; PRO0237; GPCRHHODOPSN.
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL-1; 1.
 DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL-2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
 FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 31 58 1 (POTENTIAL).
 FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 69 89 2 (POTENTIAL).
 FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 103 124 3 (POTENTIAL).
 FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 142 166 4 (POTENTIAL).
 FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 199 218 5 (POTENTIAL).
 FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 236 260 6 (POTENTIAL).
 FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 278 301 7 (POTENTIAL).
 FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 101 178 BY SIMILARITY.
 FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
 FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
 FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
 FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
 SQ SEQUENCE 352 AA; 40509 MW; 4366F148D3A5938F CRC64;

Query Match 41.3%; Score 365; DB 1; Length 352;
 Best Local Similarity 33.6%; Pred. No. 2.9e-28;
 Matches 95; Conservative 10; Mismatches 28; Indels 150; Gaps 7;

QY 20 MDYQVSSPIYDINYTSEPCQKINVKQIAA----- 49
 Db 1 MDYQVSSPTDYDYTSEPCQKINVKQIAARLLPLYSLVFIFGVGNILVVLINCKR 60
 QY 50 -----YKGLC-----AAQWDFGNTMCOHQRVHGHHSYKCG--- 83
 Db 61 LKSMTDIYLLNLAIISDLFFLLTPVFWAHYAAQWDFGNTMC--QLLTGLYPIGFSGIFF 118
 QY 84 ----- 124
 Db 119 IILLTDIDRYLAIVHAFALKARTVTGVVTSVITWVAVFASLPGIIFTRSQREGHLHYTC 178
 QY 98 SSHFPYSQYQFKNFOTLKI-----LCTRSQKEGLHYTC 97
 Db 179 SSHFPYSQYQFKNFOTLKIIVGLVPLLVNVICYSGLKTLRCRNEKKHRAVRLIF 238
 QY 125 -----GGSYKCGLC-----QEFFGLNCSNRLDGHQRV 154
 Db 239 TIMIVYFLFWAPYNIIVLLNTFOEFFGLNCSNRLDQAMQV 281

RESULT 9

CKR5 GORGO
 ID CKR5 GORGO STANDARD; PRT; 352 AA.
 AC P56439;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CCR5).
 GN CKR5 OR CMKRS.
 OS Gorilla gorilla gorilla (Lowland gorilla).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
 OX NCBI_TaxID=9595;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97268687; PubMed=9108095;
 RA Edinger A.L., Amedee A., Miller K., Doranz B.J., Endres M.,
 RA Sharron M., Samson M., Lu Z.H., Clements J.E., Murphy-Corb M.,

Peiper S.C., Parmentier M., Broder C.C., Doms R.W.; Doms R.W.;
 "Differential utilization of CCR5 by macrophage and T cell tropic
 simian immunodeficiency virus strains."
 Proc. Natl. Acad. Sci. U.S.A. 94:4005-4010(1997).
 -!- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 DIFFERENTIATION.
 -!- SUBCELLULAR LOCATION: Integral membrane protein.
 -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch)

 EMBL; AF005659; AAC62553.1; -
 InterPro; IPR000276; GPCR_Rhodopsn.
 Pfam; PF00001; 7cml.1;
 PRINTS; PR00237; GPCR_Rhodopsn.
 PROSITE; PS00237; G-PROTEIN RECP F1.1; 1.
 PROSITE; PS00262; G-PROTEIN RECP F1.2; 1.
 G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
 DOMAIN 1 30
 TRANSMEM 31 58
 DOMAIN 59 68
 TRANSMEM 69 88
 DOMAIN 90 102
 TRANSMEM 103 124
 DOMAIN 125 141
 TRANSMEM 142 166
 DOMAIN 167 198
 TRANSMEM 199 218
 DOMAIN 219 235
 TRANSMEM 236 260
 DOMAIN 261 277
 TRANSMEM 278 301
 DOMAIN 302 352
 DISULFID 101 178
 MOD_RES 3 3
 MOD_RES 10 10
 MOD_RES 14 14
 MOD_RES 15 15
 SEQUENCE 352 AA; 40515 MW; D086FCB9FE5EAC84 CRC64;

 Query Match 41.2%; Score 364; DB 1; Length 352;
 Best Local Similarity 33.9%; Pred. No. 3.7e-28;
 Matches 96; Conservative 8; Mismatches 29; Indels 150; Gaps 7;

 20 MDYQVSSPYDINVTSPCKINVKQTAA-----
 1 MDYQVSSPYDIDYTSPTCKTNVKQIAARLLPLSLVFIFGVGMVLILINCKR 60

 50 -----YKGLC-----AAQNDFGNTQHQRVGHGHHSYKCG--- 83
 61 LKSMTDIYLLNLAISDLFLTLVPFMAHVAQAQDFGNTMC--QLLTGLYIFGFGIF 118

 84 -----LCTRQKGLHYTC 97
 119 ITLLTDRLYLAIVHAFALKARTVTFGVVTSVITVWVAFSLPGIIFTRSQKGLHYTC 178

 98 SSHFPYSQYQFKNQFOTLKI-----HQRVHGG----- 124

 179 SSHFPYSQYQFKNQFOTLKIIVLGLVPLVMVICYSGLTKLLRCRNEKKRHRVRLIF 238

 125 -----GGSYKCGLC-----QEPFGLNCCSSRLDGHQVR 154
 239 TIMIVYFLFWAPNIVLLNTLTQEPFGLNCCSSRLDQAMQV 281

RESULT 10
 CCR5_MACMU
 ID CCR5_MACMU STANDARD; PRT; 352 AA.
 AC P79436; O02746;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE C-C chemokine receptor type 5 (C-C CCR-5) (CCR-5) (CCR5).
 GN CCR5 OR CCR5.
 OS Macaca mulatta (Rhesus macaque),
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), and
 OS Macaca nemestrina (Pig-tailed macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9544, 9541, 9545;
 RN [1] NCBI_TaxID=9544, 9541, 9545;
 RP SEQUENCE FROM N.A.
 RC SPECIES=M.mulatta;
 RX MEDLINE=97184592; PubMed=9032394;
 RA Marcon L., Choe H., Martin K.A., Farzan M., Ponath P.D., Wu L.,
 RA Newman W., Gerard N., Gerard C., Sodroski J.;
 RT "Utilization of C-C chemokine receptor 5 by the envelope
 glycoproteins of a pathogenic simian immunodeficiency virus."
 RT SIVmac239.";
 RL J. Virol. 71:2522-2527(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=M.mulatta; STRAIN=Indian macaque;
 RX MEDLINE=97213934; PubMed=9060623;
 RA Chen Z., Zhou P., Ho D.D., Landau N.R., Marx P.A.;
 RT "Genetically divergent strains of simian immunodeficiency virus use
 CCR5 as a coreceptor for entry."
 RT J. Virol. 71:2705-2714(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=M.mulatta;
 RX MEDLINE=21354176; PubMed=11461684;
 RA Margulies B.J., Hauer D.A., Clements J.E.;
 RT "Identification and comparison of eleven rhesus macaque chemokine
 receptors.";
 RL AIDS Res. Hum. Retroviruses 17:981-986(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES=M.mulatta, M.fascicularis, and M.nemestrina;
 RX MEDLINE=97268687; PubMed=9108095;
 RA Edinger A.L., Amedee A., Miller K., Doranz B.J., Endres M.,
 RA Sharron M., Samson M., Lu Z.H., Clements J.E., Murphy-Corb M.,
 RA Peiper S.C., Parmentier M., Broder C.C., Doms R.W.;
 RT "Differential utilization of CCR5 by macrophage and T cell tropic
 simian immunodeficiency virus strains."
 RL Proc. Natl. Acad. Sci. U.S.A. 94:4005-4010(1997).
 CC -!- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 DIFFERENTIATION.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch)

 EMBL; U77672; AAC51109.1; -
 EMBL; U73739; AAC51158.1; -
 EMBL; U96762; AAC34132.1; -
 EMBL; AF005660; AAC62554.1; -


```

61 LKSMTDIYLLNLAIISDLFLTLVTFVMAHYAAQWDFGNTWC--QLLTGLYFIFGFSGIF 118
84 -----LCTRSQKEGLHYTC 97
119 ILLTIDRYLAIVHVPALKARTVTFGVTSVITWVAVFASLPGLIFTRSQKEGLHYTC 178
98 SSHFPYSQYQFWKQFQTLKI-----HORVHG----- 124
179 SPHPYSQYQFWKQFQTLKIIVILGLVPLLVVVICYSGLKTLRLCRNKKRRAVRLIF 238
125 -----GGSYKCGLC-----QEFGLNCCSSNRLDGQHV 154
239 TIMIVYFLWAPYINIVLLNTFQEFFGLNCCSSNRLDQAMQV 281

RESULT 12
KRS_HYLLE
C CKR5_HYLLE STANDARD; PRT; 352 AA.
C 097883;
T 30-MAY-2000 (Rel. 39, Last sequence update)
T 28-FEB-2003 (Rel. 41, Last annotation update)
DE C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CCR5).
IN CCR5 OR CMKBR5
XS Hylobates leucogenys (White-cheeked gibbon).
XC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
XC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
NCBI_TaxID=61853;
IN [1]
SEQUENCE FROM N.A.
X MEDLINE=99416438; PubMed=10486970;
A Zhang Y.-W., Ryder O.A., Zhang Y.-P.;
T "Sequence evolution of the CCR5 chemokine receptor gene in primates.";
L Mol. Biol. Evol. 16:1145-1154(1999).
C -!- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
C MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
C INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
C IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
C DIFFERENTIATION.
C -!- SUBCELLULAR LOCATION: Integral membrane protein.
C -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
C
C This SWISS-PROT entry is copyright. It is produced through a collaboration
C between the Swiss Institute of Bioinformatics and the EMBL outstation -
C the European Bioinformatics Institute. There are no restrictions on its
C use by non-profit institutions as long as its content is in no way
C modified and this statement is not removed. Usage by and for commercial
C entities requires a license agreement (See http://www.isb-sib.ch/announce/
C or send an email to license@isb-sib.ch).
C
C EMBL: AF075451; AAD19863.1;
R InterPro: IPR000276; GPCR_Rhodpsn.
R Pfam: PF00001; 7cm1.1.
R PRINTS; PR00237; GPCRHHODPSN.
R PROSITE; PS00237; G-PROTEIN RECP FL1; 1.
R PROSITE; PS0262; G-PROTEIN RECP FL1.2; 1.
W G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
T DOMAIN 1 30
T TRANSMEM 31 58
T DOMAIN 59 68
T TRANSMEM 69 89
T DOMAIN 90 102
T TRANSMEM 103 124
T DOMAIN 125 141
T TRANSMEM 142 166
T DOMAIN 167 198
T TRANSMEM 199 218
T DOMAIN 219 235
T TRANSMEM 236 260
T DOMAIN 261 277
T TRANSMEM 278 301
T DOMAIN 302 352

```

```

FT DISULFID 101 178 BY SIMILARITY.
FT MOD RES 3 3 SULFATION (BY SIMILARITY).
FT MOD RES 10 10 SULFATION (BY SIMILARITY).
FT MOD RES 14 14 SULFATION (BY SIMILARITY).
SQ SEQUENCE 352 AA; 40445 MW; 4F8E4F344CEB7C91 CRC64;

Query Match 40.68; Score 359; DB 1; Length 352;
Best Local Similarity 33.94; Pred. No. 1.1e-27; Indels 150; Gaps 7;
Matches 96; Conservative 8; Mismatches 29;

QY 20 MDYQVSSPIYDINTYSEPCOKINVKQIAA----- 49
DB 1 MDYQVSSPTDYDIDYDTSEPCOKINVKQIAARLLPPLYSLVIFGFVGNMLVILVILNCR 60
QY 50 -----YKCGLC-----AAAQWDFGNTWCQHORVHGHHHSYKCG--- 83
DB 61 LKSMTDIYLLNLAIISDLFLTLVTFVMAHYAAQWDFGNTWC--QLLTGLYFIFGFSGIF 118
QY 84 -----LCTRSQKEGLHYTC 97
DB 119 ILLTIDRYLAIVHVPALKARTVTFGVTSVITWVAVFASLPGLIFTRSQKEGLHYTC 178
QY 98 SSHFPYSQYQFWKQFQTLKI-----HORVHG----- 124
DB 179 SPHPYSQYQFWKQFQTLKIIVILGLVPLLVVVICYSGLKTLRLCRNKKRRAVRLIF 238
QY 125 -----GGSYKCGLC-----QEFGLNCCSSNRLDGQHV 154
DB 239 TIMIVYFLWAPYINIVLLNTFQEFFGLNCCSSNRLDQAMQV 281

RESULT 13
CKR5_CERAE
ID CKR5_CERAE STANDARD; PRT; 352 AA.
AC P56493;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (CCR5).
GN CCR5 OR CMKBR5
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]_TaxID=9534;
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=98001387; PubMed=9343222;
RA Kuhnmann S.E., Platt E.J., Kozak S.L., Kabat D.;
RT "Polymorphisms in the CCR5 genes of African green monkeys and mice
RT implicate specific amino acids in infections by simian and human
RT immunodeficiency viruses.";
RL J. Virol. 71:8642-8656(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Murayama Y., Matsunaga S., Inoue-Murayama M.;
RT "cDNA sequence of African green monkey CCR-5 chemokine receptor
RT gene.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
CC DIFFERENTIATION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

DR EMBL; U83324; AAC51795.1; -
DR EMBL; U83325; AAC51796.1; -
DR EMBL; AB015944; BAA31328.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G-PROTEIN RECP_F1_1; 1.
DR PROSITE; PS00262; G-PROTEIN RECP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation;
KW Polymorphism.
FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 31 58 1 (POTENTIAL).
FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 69 89 2 (POTENTIAL).
FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 103 124 3 (POTENTIAL).
FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 142 166 4 (POTENTIAL).
FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 199 218 5 (POTENTIAL).
FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 236 260 6 (POTENTIAL).
FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 278 301 7 (POTENTIAL).
FT DOMAIN 302 352 BY SIMILARITY.
FT DISULFID 101 178 BY SIMILARITY.
FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
FT VARIANT 14 14 N -> Y.
FT VARIANT 352 352 F -> L.
SQ SEQUENCE 352 AA; 40561 MW; 7F52E690C72EC29A CRC64;
Query Match 40.4%; Score 357; DB 1; Length 352;
Best Local Similarity 33.6%; Pred. No. 1.7e-27;
Matches 95; Conservative 9; Mismatches 29; Indels 150; Gaps 7;
QY 20 MDYQSSPYDINVTSEPCOKINVKQIAA----- 49
Db 1 MDYQSSPYDINVTSEPCOKINVKQIAARLLPLYSLVIFQVGNILVLLINCKR 60
QY 50 -----YKCGLC-----AAQWDFGNTMCQHQRVHGHHSYKCG--- 83
Db 61 LKSMTDIVLLNLAISDLLFLTLVTFPMAYAAQWDFGNTMC--QLLTGLYIFGFSGIFF 118
QY 84 ----- 178
Db 119 ILLTIDRYLAIVHAFVAFKARTVTFGVVTSVITVWVAVFASLPRIIFTRSQREGLHYTC 178
QY 98 SSHFPYSQVQFNQFOTLKI-----HORVHGG----- 124
Db 179 SSHFPYSQVQFNQFOTLKIIVILGLVPLLVWVICYSGILKTLRCNEKKRRAVRLIF 230
QY 125 -----GSGYKCGLC-----QEFFGLNCCSSNRLDGHQV 154
Db 239 TIMIVYFLFWAPYINVLNLTTFQEFFGLNCCSSNRLDQAMQV 281

```

RESULT 14
CKR5 MOUSE

ID CKR5 MOUSE STANDARD; PRT; 354 AA.
AC P51682; O35313; O35891; P97308; P97405; Q61867;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (MIP-1
alpha receptor).
GN CKR5 OR CMKBR5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ; TISSUE=Spleen;
RX MEDLINE=96205938; PubMed=8631787;
RA Boring L., Gosling J., Montecarlo F.S., Lusis A.J., Tsou C.-L.,
RA Charo I.F.;
RT "Molecular cloning and functional expression of murine JE (monocyte
RT chemoattractant protein 1) and murine macrophage inflammatory protein
RT 1alpha receptors: evidence for two closely linked C-C chemokine
RT receptors on chromosome 9."
RL J. Biol. Chem. 271:7551-7558(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 X CBA; TISSUE=Thymus;
RX MEDLINE=96278910; PubMed=8662890;
RA Meyer A., Coyle A.J., Proudfoot A.E.I., Wells T.N.C., Power C.A.;
RT "Cloning and characterization of a novel murine macrophage
RT inflammatory protein-1 alpha receptor."
RL J. Biol. Chem. 271:14445-14451(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Ola;
RX Kuziel W.A., Beck M.A., Dawson T.C., Maeda N.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6, and NIH Swiss; TISSUE=Kidney, Liver, and Spleen;
RX MEDLINE=98001387; PubMed=9343222;
RA Kuhlmann S.E., Platt B.J., Kozak S.L., Kabat D.;
RT "Polymorphisms in the CCR5 genes of African green monkeys and mice
RT implicate specific amino acids in infections by simian and human
RT immunodeficiency viruses."
RL J. Virol. 71:8642-8656(1997).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RX MEDLINE=97404635; PubMed=9261347;
RA Doranz B.J., Lu Z.H., Rucker J., Zhang T.Y., Sharron M., Cen Y.H.,
RA Wang Z.X., Guo H.H., Du J.G., Accavitti M.A., Doms R.W., Peiper S.C.;
RT "Two distinct CCR5 domains can mediate coreceptor usage by human
RT immunodeficiency virus type 1."
RL J. Virol. 71:6305-6314(1997).
RN [6]
RP SEQUENCE FROM N.A.
RC Guo B., Kuno K., Harada A., Matsushima K.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: DETECTED IN MONOCYTE/MACROPHAGE CELL LINES,
CC BUT NOT IN NONHEMATOPOIETIC CELL LINES.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

CC EMBL; U47036; AAC52454.1; -
CC EMBL; X34151; CAA63867.1; -
CC EMBL; U68565; AAB37273.1; -
CC EMBL; U83327; AAC53386.1; -
CC EMBL; AF022990; AAC53389.1; -
CC EMBL; AF019772; AAB71183.1; -
CC EMBL; D83648; BAA12024.1; -
CC MGDI; MG1:107182; Ccr5.
GO; GO:0016493; F1C-C chemokine receptor activity; IDA.

GO: GO:0006952; P: defense response; IMP.
InterPro: IPR000276; GPCR_Rhodpsn.
Pfam: PF00001; 7tm.1; 1.

PRINTS: PR00237; GPCRHHODPSN.
PROSITE: PS00237; G-PROTEIN RECEPTOR FL 1; 1.
PROSITE: PS00237; G-PROTEIN RECEPTOR FL 2; 1.

PROSITE: PS00262; G-PROTEIN RECEPTOR FL 1; 1.
G-protein coupled receptor; Transmembrane; Polymorphism.
DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

DOMAIN 1 32
EXTRACELLULAR (POTENTIAL).

Db 1 MDPQSGIPTIYIDYDYSMSAPCKVNVKQIAAQLLPPLYSLVFIFGPGNNMVFILILISC 60
QY 50 -----YKGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG- 83
Db 61 KXLSMTDIYLFNLAI S D L L F L L T L P F W A H Y A A N E W F G N I M C --K L F T G I Y H I G Y F G G I 118
QY 84 -----LCTRSOKEG LHY 95
Db 119 PFIILLTIDRYLAIVHAVFAIKARTVNGVITSVTVWVAVFVSLPEIIFMRSQKESHY 178
QY 96 TCSSHPYQYQFQWKNFQTLK-----IHQRVHGG--GGSYKC----- 130
Db 179 TCSPFELHIQYRFWKGH F Q L K W V I L S L I L P L L V M V I C Y S G I L N T L F R C N E K K R H R A V R L 238
QY 131 -----GLCQEFGLNCCSSNRLDGHQV 154
Db 239 IFAIMIVFELFWTPYINIVILLTTFOEYFGLNCCSSNRLDQAMQV 283

Search completed: October 21, 2003, 17:51:31
Job time : 16.9017 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 17:47:34 : Search time 64.9288 Seconds
(without alignments)
623.980 Million cell updates/sec

Title: US-10-057-890A-31

Perfect score: 884

Sequence: 1 MKVSVAAALSCLMLVTALGSM.....GLNNCSSNSRLDGHQHVHAA 157

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	373	42.2	352	018772	018772 pan troglod
2	370	41.9	344	09TOR8	09tor8 cercocobus
3	370	41.9	344	077833	077833 cercocobus
4	369	41.7	352	095NC5	095nc5 hylobates s
5	369	41.7	352	095NC3	095nc3 miopithecus
6	367	41.5	352	095NC3	095nc3 miopithecus
7	367	41.5	352	018771	018771 pan troglod
8	366	41.4	352	09TSK1	09tsk1 cercopithec
9	366	41.4	352	09TW49	09tw49 cercocobus
10	366	41.4	352	09XT13	09xt13 papio anubi
11	366	41.4	352	09TW45	09tw45 cercopithec
12	366	41.4	352	095NE8	095ne8 cercopithec
13	365	41.3	352	09XT14	09xt14 colobus gue
14	365	41.3	352	095NC6	095nc6 trachypithe
15	365	41.3	352	095NC8	095nc8 colobus pol
16	364	41.2	352	095NC9	095nc9 gorilla gor

17	363	41.1	352	6	097975	097975 macaca arct
18	363	41.1	352	6	09MZA2	09mza2 cercopithec
19	362	41.0	352	6	095NC1	095nc1 theropithec
20	362	41.0	352	6	095NC0	095nc0 hylobates m
21	362	41.0	352	6	09TV46	09tv46 cercopithec
22	362	41.0	352	6	097962	097962 pygathrix a
23	361	40.8	352	6	095NC7	095nc7 naalalis lar
24	361	40.8	352	6	09TV42	09tv42 cercopithec
25	361	40.8	352	6	077776	077776 cercocobus
26	361	40.8	352	6	095NE1	095ne1 cercocobus
27	361	40.8	352	6	09TQX0	09tqx0 cercopithec
28	359	40.6	352	6	09TV43	09tv43 cercopithec
29	358	40.5	352	6	018770	018770 pan troglod
30	357	40.4	352	6	09TV47	09tv47 cercopithec
31	357	40.4	352	6	09BGN5	09bgn5 cercopithec
32	357	40.4	352	6	09XT12	09xt12 cercopithec
33	357	40.4	352	6	09TSQ7	09tsq7 cercopithec
34	356	40.3	352	6	095ND2	095nd2 mandrillus
35	356	40.3	352	6	095ND1	095nd1 mandrillus
36	355	40.2	352	6	09XT76	09xt76 cercopithec
37	355	40.2	352	6	09TV44	09tv44 cercopithec
38	355	40.2	352	6	09XS35	09xs35 macaca neme
39	355	40.2	352	6	09MZA3	09mza3 hylobates a
40	355	40.2	352	6	09TV93	09tv93 macaca arct
41	355	40.2	352	6	095ND0	095nd0 erythrocebu
42	346	39.1	352	6	09BGN6	09bgn6 cercopithec
43	345	39.0	339	4	09UN25	09un25 homo sapien
44	344	38.9	339	4	09UN23	09un23 homo sapien
45	344	38.9	339	4	09UN27	09un27 homo sapien

ALIGNMENTS

RESULT 1

018772 018772 PRELIMINARY; PRT; 352 AA.
AC 018772;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE CCR5 receptor (Fragment).
GN CCR5.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
OX NCBI_TaxID:9598;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHCCRS-142a;
RA Zhang L., Carruthers C.D., He T., Huang Y., Cao Y., Wang G., Hahn B.,
RA Ho D.D.;
RT "HIV-1 subtypes, co-receptor usage, and CCR5 polymorphism."
RL AIDS Res. Hum. Retroviruses 0:0-0(1997).
DR EMBL; AF011541; AAB65741.1; -;
DR InterPro; IPR000276; GPCR_Rhodop.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; G PROTEIN RECP F1.1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECP_F1_2; 1.
KW Receptor.
FT NON TER 352 352
SQ SEQUENCE 352 AA; 4059 MW; 39688FA7004C952P CRC64;

Query Match 42.2%; Score 373; DB 6; Length 352;

Best Local Similarity 34.8%; Pred.No. 1.8e-33;

Matches 98; Conservative 8; Mismatches 27; Indels 150; Gaps 7;

QY 20 MDYQSSPIYDINYTSEPCQKINVKQIAA-----49

Db 1 MDYQSSPIYDINYTSEPCQKINVKQIAA-----49

QY 50 -----YKGLC-----AAQWDFGNTWCQHQRVHGHHSYKCG---83


```

RA Diop O., Rigoulet J., Barre-Sinoussi F., Fomsgaard A.;
RT "Mutations in CCR5-coding sequences are not associated with HIV
RL carrier status in African nonhuman primates.";
RL AIDS Res. Hum. Retroviruses 15:931-939(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=1410;
RA Mueller-Trutwin M.C., Corbet S., Hansen J., Diop O.,
RA Georges-Courbot M.-C., Barre-Sinoussi F., Fomsgaard A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF035214; AAD44007.1; -.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
DR Receptor.
KW Receptor.
SQ SEQUENCE 352 AA; 40481 MW; 2578A0F2C07A4A65 CRC64;

Query Match 41.7%; Score 369; DB 6; Length 352;
Best Local Similarity 34.3%; Pred. No. 5e-33;
Matches 97; Conservative 8; Mismatches 28; Indels 150; Gaps 7;

QY 20 MDYQSSPIYDINYYTSEPCQKINVKQIAA----- 49
DB 1 MDYQSSPTDYDYTSEPCQKINVKQIAAARLLPPLYSLVFIFGVGNILVVLINCKR 60
QY 50 -----YKGLC-----AAQWDFGNTMCOHQRVHGHHSYKCG--- 83
DB 61 LKSMTDIYLLNLAISDLFLLTVPFWAHYAAQWDFGNTMC--QLLTGLYFIFGSGIFF 118
QY 84 ----- 124
DB 119 IILLTDIYLAIVHAFVALKARTVFGVTVSVITWVAVFASLPGIIFTRSQEGLHYTC 178
QY 98 SSHFPYSQYQFQWNPOTLKI-----HORVHG----- 124
DB 179 SSHFPYSQYQFQWNPOTLKIIVLGLVPLLVVICYSGLIKTLRCRNEKGRHRAVLIF 238
QY 125 -----GSGYKGLC-----QEFFGLNCCSSNRLDGHQV 154
DB 239 TIMIVVFLWPAPYNIIVLLNTFQEFFGLNCCSSNRLDQMQV 281

RESULT 6
Q95NC3 ID Q95NC3 PRELIMINARY; PRT; 352 AA.
AC Q95NC3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE C-C chemokine receptor 5.
GN CCR5.
OS Miopithecus talapoin (Talapoin) (Cercopithecus talapoin).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Miopithecus.
OX NCBI_TaxID=36231;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang Y., Ryder O.A., Zhang Y.;
RT "Sequence comparison of the CCR5 gene in primates and primate
RT phylogeny.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF177886; AAK43369.1; -.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
DR Receptor.
KW Receptor.
SQ SEQUENCE 352 AA; 40546 MW; 6464152F3B566A85 CRC64;

RA Diop O., Rigoulet J., Barre-Sinoussi F., Fomsgaard A.;
RT "Mutations in CCR5-coding sequences are not associated with HIV
RL carrier status in African nonhuman primates.";
RL AIDS Res. Hum. Retroviruses 15:931-939(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=1410;
RA Mueller-Trutwin M.C., Corbet S., Hansen J., Diop O.,
RA Georges-Courbot M.-C., Barre-Sinoussi F., Fomsgaard A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF035214; AAD44007.1; -.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
DR Receptor.
KW Receptor.
SQ SEQUENCE 352 AA; 40481 MW; 2578A0F2C07A4A65 CRC64;

Query Match 41.7%; Score 369; DB 6; Length 352;
Best Local Similarity 34.3%; Pred. No. 5e-33;
Matches 97; Conservative 8; Mismatches 28; Indels 150; Gaps 7;

QY 20 MDYQSSPIYDINYYTSEPCQKINVKQIAA----- 49
DB 1 MDYQSSPTDYDYTSEPCQKINVKQIAAARLLPPLYSLVFIFGVGNILVVLINCKR 60
QY 50 -----YKGLC-----AAQWDFGNTMCOHQRVHGHHSYKCG--- 83
DB 61 LKSMTDIYLLNLAISDLFLLTVPFWAHYAAQWDFGNTMC--QLLTGLYFIFGSGIFF 118
QY 84 ----- 124
DB 119 IILLTDIYLAIVHAFVALKARTVFGVTVSVITWVAVFASLPGIIFTRSQEGLHYTC 178
QY 98 SSHFPYSQYQFQWNPOTLKI-----HORVHG----- 124
DB 179 SSHFPYSQYQFQWNPOTLKIIVLGLVPLLVVICYSGLIKTLRCRNEKGRHRAVLIF 238
QY 125 -----GSGYKGLC-----QEFFGLNCCSSNRLDGHQV 154
DB 239 TIMIVVFLWPAPYNIIVLLNTFQEFFGLNCCSSNRLDQMQV 281

RESULT 5
Q9TV50 ID Q9TV50 PRELIMINARY; PRT; 352 AA.
AC Q9TV50;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CC chemokine receptor type 5.
GN CCR5.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Pan.
OX NCBI_TaxID=9599;
RN [1]
RP SEQUENCE FROM N.A.
RA Muller-Trutwin M.C., Corbet S., Hansen J., Georges-Courbot M.C.,
RX MEDLINE=99335215; PubMed=10408730;
RC STRAIN=1410;
RL Muller-Trutwin M.C., Corbet S., Hansen J., Georges-Courbot M.C.,

```



```

Query Match      41.5%; Score 367; DB 6; Length 352;
Best Local Similarity 33.9%; Pred. No. 8.4e-33;
Matches 96; Conservative 9; Mismatches 28; Indels 150; Gaps 7;

2Y 20 MDYQSSPIYDINYYTSEPCQKINVKQIAA----- 49
DB 1 MDYQSSPIYDINYYTSEPCQKINVKQIAARLLPLYSLVPIFGVGNILVLLINCKR 60
2Y 50 -----YKGLC-----AAQWDFGNTMCHQORVGHGHHHSYKCG--- 83
DB 61 LKSMTDIYLLNLAISDLFLLTVPFWAHYAAQWDFGNTMC--RLTLGLYTFIGFSGIFF 118
2Y 84 -----YKGLC-----LCTRSQKGLHYTC 97
DB 119 ILLTIDRYLAIVHAFVALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKGLHYTC 178
2Y 98 SSHFPYSQYQFWKFNQTLKI-----HORVHG----- 124
DB 179 SSHFPYSQYQFWKFNQTLKIVILGLVPLLVVICYSGLKTLRCRNEKKRHRAVLIF 238
2Y 125 -----GGSYKGLC-----QEFFGLNCCSSNRLDGHQV 154
DB 239 TIMIVYFLFWAPYINVLNTTQEFFGLNCCSSNRLDQAMQV 281

RESULT 7
ID O18771 PRELIMINARY; PRT; 352 AA.
AC O18771;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE CCR5 receptor (Fragment).
EN CCR5.
NC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
XC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Pan.
XX NCBI_TaxID=9598;
IN [1]
IP SEQUENCE FROM N.A.
IC STRAIN=ChCERS-141a;
IA Ho D.D.;
JA Zhang L., Carruthers C.D., He T., Huang Y., Cao Y., Wang G., Hahn B.,
T "HIV-1 subtypes, co-receptor usage, and CCR5 polymorphism.";
TL AIDS Res. Hum. Retroviruses 0:0-0(1997).
DR EMBL; AF011539; AAB65739.1;
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
CW Receptor.
T NON_TER
T SEQUENCE 352 352
T SEQUENCE 352 AA; 40466 MW; 3FFAC7ABAE1D4FB CRC64;

Query Match      41.5%; Score 367; DB 6; Length 352;
Best Local Similarity 34.3%; Pred. No. 8.4e-33;
Matches 97; Conservative 8; Mismatches 28; Indels 150; Gaps 7;

2Y 20 MDYQSSPIYDINYYTSEPCQKINVKQIAA----- 49
DB 1 MDYQSSPIYDINYYTSEPCQKINVKQIAARLLPLYSLVPIFGVGNILVLLINCKR 60
2Y 50 -----YKGLC-----AAQWDFGNTMCHQORVGHGHHHSYKCG--- 83
DB 61 LKSMTDIYLLNLAISDLFLLTVPFWAHYAAQWDFGNTMC--RLTLGLYTFIGFSGIFF 118
2Y 84 -----YKGLC-----LCTRSQKGLHYTC 97
DB 119 ILLTIDRYLAIVHAFVALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKGLHYTC 178
2Y 98 SSHFPYSQYQFWKFNQTLKI-----HORVHG----- 124
DB 179 SSHFPYSQYQFWKFNQTLKIVILGLVPLLVVICYSGLKTLRCRNEKKRHRAVLIF 238
2Y 125 -----GGSYKGLC-----QEFFGLNCCSSNRLDGHQV 154
DB 239 TIMIVYFLFWAPYINVLNTTQEFFGLNCCSSNRLDQAMQV 281

```

```

DB 179 SSHFPYSQYQFWKFNQTLKIVILGLVPLLVVICYSGLKTLRCRNEKKRHRAVLIF 238
QY 125 -----GGSYKGLC-----QEFFGLNCCSSNRLDGHQV 154
DB 239 TIMIVYFLFWAPYINVLNTTQEFFGLNCCSSNRLDQAMQV 281

RESULT 8
ID Q9TSK1 PRELIMINARY; PRT; 352 AA.
AC Q9TSK1;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE G-protein coupled chemokine receptor. (Grivet).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RA Holtkamp N., Baier M., Werner A.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF019379; AAD01639.1;
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
CW Receptor.
XW Receptor.
SQ SEQUENCE 352 AA; 40588 MW; 0F1869D9A6668DBB CRC64;

Query Match      41.4%; Score 366; DB 6; Length 352;
Best Local Similarity 33.9%; Pred. No. 1.1e-32;
Matches 96; Conservative 9; Mismatches 28; Indels 150; Gaps 7;

QY 20 MDYQSSPIYDINYYTSEPCQKINVKQIAA----- 49
DB 1 MDYQSSPIYDINYYTSEPCQKINVKQIAARLLPLYSLVPIFGVGNILVLLINCKR 60
QY 50 -----YKGLC-----AAQWDFGNTMCHQORVGHGHHHSYKCG--- 83
DB 61 LKSMTDIYLLNLAISDLFLLTVPFWAHYAAQWDFGNTMC--QLTLGLYTFIGFSGIFF 118
QY 84 -----LCTRSQKGLHYTC 97
DB 119 ILLTIDRYLAIVHAFVALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKGLHYTC 178
QY 98 SSHFPYSQYQFWKFNQTLKI-----HORVHG----- 124
DB 179 SSHFPYSQYQFWKFNQTLKIVILGLVPLLVVICYSGLKTLRCRNEKKRHRAVLIF 238
QY 125 -----GGSYKGLC-----QEFFGLNCCSSNRLDGHQV 154
DB 239 TIMIVYFLFWAPYINVLNTTQEFFGLNCCSSNRLDQAMQV 281

RESULT 9
ID Q9TV49 PRELIMINARY; PRT; 352 AA.
AC Q9TV49;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE CC chemokine receptor type 5 (C-C chemokine receptor 5).
GN CCR5.
OS Cercocebus galeries (Agile mangabey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Cercocebus.
OX NCBI_TaxID=9532;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=4;
RX MEDLINE=99335215; PubMed=10408730;
RA Muller-Trutwin M.C., Corbet S., Hansen J., Georges-Courbot M.C.,
RA Diop O., Rigoulet J., Barre-Sinoussi F., Fomsgaard A.;
RT "Mutations in CCR5-coding sequences are not associated with SIV
RT carrier status in African nonhuman primates.";
RL AIDS Res. Hum. Retroviruses 15:931-939 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=4;
RA Mueller-Trutwin M.C., Corbet S., Hansen J., Diop O.,
RA Georges-Courbot M.C., Barre-Sinoussi F., Fomsgaard A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Zhang Y., Ryder O.A., Zhang Y.;
RT "Sequence comparison of the CCR5 gene in primates and primate
RT phylogeny.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF035215; AAD44008.1; -
DR EMBL; AF177898; AAK43381.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 352 AA; 40479 MW; 5E1504A9B87278B2 CRC64;

Query Match 41.4%; Score 366; DB 6; Length 352;
Best Local Similarity 33.9%; Pred. No. 1.1e-32;
Matches 96; Conservative 9; Mismatches 28; Indels 150; Gaps 7;

Qy 20 MDYQVSSPYDINYTSEPCQKINVKQIAA-----49
Db 1 MDYQVSSPYDIDYITSEPCQKINVKQIAARLLPLYSLVIFGVGNILVVLINCKR 60

Qy 50 -----YKCGLC-----AAQWDFGNTMCOHORVGHGHHHSYKCG---83
Db 61 LKSMTDIYLLNLAISDLLFLLTVPFWAHYAAQWDFGNTMC--QLLTGLYFIGFPGSGLFF 118

Qy 84 -----LCTRSQKEGLHYTC 97
Db 119 IILLTIDRYLAIVHAFVALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQEGLHYTC 178

Qy 98 SSHFPYSQYQFQWNTQTLKI-----HQRVHGG-----124
Db 179 SSHFPYSQYQFQWNTQTLKIVILGLVPLLVWVICYSGILKTLRCRNEKRRHRAVRLIF 238

Qy 125 -----GGSYKCGLC-----QEFFGLNCCSSNRLDGHQRV 154
Db 239 TIMIVFLFWAPYNIIVLLNTFOEFFGLNCCSSNRLDQAMQV 281

RESULT 10
Q9XT13 PRELIMINARY; PRT; 352 AA.
AC Q9XT13;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CC Chemokine receptor 5.
GN CCR5.
OS Papio anubis (Olive baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Papio.
OX NCBI_TaxID=9555;
RN [1]
RP SEQUENCE FROM N.A.
RA Spearman P.W., Mburu D.N., Graham B.S.;
RT "Differential Utilization of CCR5 Molecules from Three East African

```

```

RT Simian Species by the HIV-1 Envelope Glycoprotein.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF141640; AAD32685.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 352 AA; 40489 MW; 0B47E337C11E2E1E CRC64;

Query Match 41.4%; Score 366; DB 6; Length 352;
Best Local Similarity 33.9%; Pred. No. 1.1e-32;
Matches 96; Conservative 9; Mismatches 28; Indels 150; Gaps 7;

Qy 20 MDYQVSSPYDINYTSEPCQKINVKQIAA-----49
Db 1 MDYQVSSPYDIDYITSEPCQKINVKQIAARLLPLYSLVIFGVGNILVVLINCKR 60

Qy 50 -----YKCGLC-----AAQWDFGNTMCOHORVGHGHHHSYKCG---83
Db 61 LKSMTDIYLLNLAISDLLFLLTVPFWAHYAAQWDFGNTMC--QLLTGLYFIGFPGSGLFF 118

Qy 84 -----LCTRSQKEGLHYTC 97
Db 119 IILLTIDRYLAIVHAFVALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQEGLHYTC 178

Qy 98 SSHFPYSQYQFQWNTQTLKI-----HQRVHGG-----124
Db 179 SSHFPYSQYQFQWNTQTLKIVILGLVPLLVWVICYSGILKTLRCRNEKRRHRAVRLIF 238

Qy 125 -----GGSYKCGLC-----QEFFGLNCCSSNRLDGHQRV 154
Db 239 TIMIVFLFWAPYNIIVLLNTFOEFFGLNCCSSNRLDQAMQV 281

RESULT 11
Q9TV45 PRELIMINARY; PRT; 352 AA.
AC Q9TV45;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CC Chemokine receptor type 5.
GN CCR5.
OS Cercopithecus nictitans (white-nosed guenon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Cercopithecus.
OX NCBI_TaxID=36228;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=7;
RX MEDLINE=99335215; PubMed=10408730;
RA Muller-Trutwin M.C., Corbet S., Hansen J., Georges-Courbot M.C.,
RA Diop O., Rigoulet J., Barre-Sinoussi F., Fomsgaard A.;
RT "Mutations in CCR5-coding sequences are not associated with SIV
RT carrier status in African nonhuman primates.";
RL AIDS Res. Hum. Retroviruses 15:931-939 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=7;
RA Mueller-Trutwin M.C., Corbet S., Hansen J., Diop O.,
RA Georges-Courbot M.C., Barre-Sinoussi F., Fomsgaard A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF035219; AAD44012.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 352 AA; 40658 MW; BC665891ED5661F8 CRC64;

```

```
Query Match 41.4%; Score 366; DB 6; Length 352;
Best Local Similarity 33.9%; Pred. No. 1.1e-32;
Matches 96; Conservative 9; Mismatches 28; Indels 150; Gaps 7;

20 MDYQVSSPIYDINVTSEPCQKINVKQIAA-----
1 MDYQVSSPTDIDYITSEPCQKINVKQIAARLLPLLSLVFIFGVGNILVVLINCKR 60
50 -----YKGLC-----AAQWDFGNTMCHQHVHGHHSYKCG--- 83
61 LKSMTDIYLLNLAISSLFLTLVTFWPAHYAAQWDFGNTMC--QLLTGLYFIFGSGIFF 118
84 -----LCTRSQKGLHYTC 97
119 ILLTIDRYLAIVHAFVALKARTVTFGVVTSVITWVAVFASLPRIIFTRSQREGLHYTC 178
98 SSHFPYSQYQFWKNFQTLKI-----HORVHG----- 124
179 SSHFPYSQYQFWKNFQTLKIIVLGLVPLLVWVICYSGLKTLRCRNEKKRRAVRLIF 238
125 -----GGSYKGLC-----QEFFGLNCCSSNRLDGHQV 154
239 TIMIVYFLWAPYINIVLLNTFQEFFGLNCCSSNRLDQAMQV 281

RESULT 12
Q95NE8 PRELIMINARY; PRT; 352 AA.
ID Q95NE8
AC Q95NE8
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE CC chemokine receptor type 5.
GN CC chemokine receptor type 5.
OS Cercopithecus tantalus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciidae; Cercopitheciidae.
OC Colobus.
OX NCBI_TaxID=60712;
RN [1]
RP SEQUENCE FROM N.A.
RA Muller-Trutwin M.C., Corbet S., Hansen J., Georges-Courbot M.C.,
RT "Mutations in CCR5-coding sequences are not associated with HIV
RL carrier status in African nonhuman primates."
DR AIDS Res. Hum. Retroviruses 15:931-939(1999).
DR EMBL; AF081577; AAD45495.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm1.1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsin.
DR PROSITE; PS00237; G PROTEIN RECEPTOR FL1; 1.
DR PROSITE; PS0262; G-PROTEIN RECEPTOR FL2; 1.
DR PROSITE; PS0262; G-PROTEIN RECEPTOR FL2; 1.
RN Receptor.
SQ SEQUENCE 352 AA; 40576 MW; 98789D9B7968DB CRC64;

Query Match 41.4%; Score 366; DB 6; Length 352;
Best Local Similarity 33.9%; Pred. No. 1.1e-32;
Matches 96; Conservative 9; Mismatches 28; Indels 150; Gaps 7;

20 MDYQVSSPIYDINVTSEPCQKINVKQIAA-----
1 MDYQVSSPTDIDYITSEPCQKINVKQIAARLLPLLSLVFIFGVGNILVVLINCKR 60
50 -----YKGLC-----AAQWDFGNTMCHQHVHGHHSYKCG--- 83
61 LKSMTDIYLLNLAISSLFLTLVTFWPAHYAAQWDFGNTMC--QLLTGLYFIFGSGIFF 118
84 -----LCTRSQKGLHYTC 97
119 ILLTIDRYLAIVHAFVALKARTVTFGVVTSVITWVAVFASLPRIIFTRSQREGLHYTC 178
98 SSHFPYSQYQFWKNFQTLKI-----HORVHG----- 124
179 SSHFPYSQYQFWKNFQTLKIIVLGLVPLLVWVICYSGLKTLRCRNEKKRRAVRLIF 238
125 -----GGSYKGLC-----QEFFGLNCCSSNRLDGHQV 154
239 TIMIVYFLWAPYINIVLLNTFQEFFGLNCCSSNRLDQAMQV 281

RESULT 13
Q9XT14 PRELIMINARY; PRT; 352 AA.
ID Q9XT14
AC Q9XT14
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE CC chemokine receptor 5.
GN CC chemokine receptor 5.
OS Colobus guereza (Black-and-white colobus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae; Colobinae;
OC Colobus.
OX NCBI_TaxID=33548;
RN [1]
RP SEQUENCE FROM N.A.
RA Spearman P.W., Mburu D.N., Graham B.S.;
RT "Differential Utilization of CCR5 Molecules from Three East African
RT Simian Species by the HIV-1 Envelope Glycoprotein."
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF141639; AAD32684.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm1.1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsin.
DR PROSITE; PS00237; G PROTEIN RECEPTOR FL1; 1.
DR PROSITE; PS0262; G-PROTEIN RECEPTOR FL2; 1.
DR PROSITE; PS0262; G-PROTEIN RECEPTOR FL2; 1.
RN Receptor.
SQ SEQUENCE 352 AA; 40550 MW; 9B078EF04D34DB36 CRC64;

Query Match 41.3%; Score 365; DB 6; Length 352;
Best Local Similarity 33.6%; Pred. No. 1.4e-32;
Matches 95; Conservative 10; Mismatches 28; Indels 150; Gaps 7;

20 MDYQVSSPIYDINVTSEPCQKINVKQIAA-----
1 MDYQVSSPTDIDYITSEPCQKINVKQIAARLLPLLSLVFIFGVGNILVVLINCKR 60
50 -----YKGLC-----AAQWDFGNTMCHQHVHGHHSYKCG--- 83
61 LKSMTDIYLLNLAISSLFLTLVTFWPAHYAAQWDFGNTMC--QLLTGLYFIFGSGIFF 118
84 -----LCTRSQKGLHYTC 97
119 ILLTIDRYLAIVHAFVALKARTVTFGVVTSVITWVAVFASLPRIIFTRSQREGLHYTC 178
98 SSHFPYSQYQFWKNFQTLKI-----HORVHG----- 124
179 SSHFPYSQYQFWKNFQTLKIIVLGLVPLLVWVICYSGLKTLRCRNEKKRRAVRLIF 238
125 -----GGSYKGLC-----QEFFGLNCCSSNRLDGHQV 154
239 TIMIVYFLWAPYINIVLLNTFQEFFGLNCCSSNRLDQAMQV 281

RESULT 14
Q95NC6 PRELIMINARY; PRT; 352 AA.
ID Q95NC6
AC Q95NC6
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE C-C chemokine receptor 5.
GN CCR5.
OS Trachypithecus johnii (hooded leaf monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae; Colobinae;
OC Trachypithecus.
OX NCBI_TaxID=66063;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang Y., Ryder O.A., Zhang Y.;
RT "Sequence comparison of the CCR5 gene in primates and primate
RT phylogeny.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF177883; AAK43366.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 352 AA; 4062 MW; 52824E032559F7F CRC64;

Query Match 41.3%; Score 365; DB 6; Length 352;
Best Local Similarity 33.6%; Pred. No. 1.4e-32;
Matches 95; Conservative 10; Mismatches 28; Indels 150; Gaps 7;

QY 20 MDYQVSSPIYDINYTSEPCQKINVKQIAA----- 49
Db 1 MDYQVSSPTDIDYTTSEPCQKVNKQIAARLLPPLYSLVFIFGVGNILVVLINCKR 60
QY 50 -----YKCGLC-----AAQWDFGNTMCOHORVGHGHHHSYKCG--- 83
Db 61 LKSMTDIYLLNLAIISDLFELLTPVFWAHYAAQWDFGNTMC--QLTGLYIFGFSGIFP 118
QY 84 ----- 124
Db 119 IILLTIDRYLAIVHAVFALKARTVTGVVTSVITWVAVFASLPGIIFTRSQEGLHYTC 178
QY 98 SSHFPYSQYQFWKNFOTLKI-----LCTRSQKGLHYTC 97
Db 179 SSHFPYSQYQFWKNFOTLKIIVILGLVLPVLMVICYSGILKTLRCRSEKGRHRAVLIF 238
QY 125 -----GGSYKCGLC-----QEFFGLNCCSSNRLDGHORV 154
Db 239 TIMIVYFLFWAPYINVLNLTQEFFGLNCCSSNRLDQAMQV 281

RESULT 15
Q95NC8 PRELIMINARY; PRT; 352 AA.
AC Q95NC8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE C-C chemokine receptor 5.
GN CCR5.
OS Colobus polykomos.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae; Colobinae;
OC Colobus.
OX NCBI_TaxID=9572;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang Y., Ryder O.A., Zhang Y.;
RT "Sequence comparison of the CCR5 gene in primates and primate
RT phylogeny.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF177881; AAK43364.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 352 AA; 40578 MW; 4366F149C3B4938F CRC64;

Query Match 41.3%; Score 365; DB 6; Length 352;

```

```

Best Local Similarity 33.6%; Pred. No. 1.4e-32;
Matches 95; Conservative 10; Mismatches 28; Indels 150; Gaps 7;

QY 20 MDYQVSSPIYDINYTSEPCQKINVKQIAA----- 49
Db 1 MDYQVSSPTDIDYTTSEPCQKVNKQIAARLLPPLYSLVFIFGVGNILVVLINCKR 60
QY 50 -----YKCGLC-----AAQWDFGNTMCOHORVGHGHHHSYKCG--- 83
Db 61 LKSMTDIYLLNLAIISDLFELLTPVFWAHYAAQWDFGNTMC--QLTGLYIFGFSGIFP 118
QY 84 ----- 124
Db 119 IILLTIDRYLAIVHAVFALKARTVTGVVTSVITWVAVFASLPGIIFTRSQEGLHYTC 178
QY 98 SSHFPYSQYQFWKNFOTLKI-----LCTRSQKGLHYTC 97
Db 179 SSHFPYSQYQFWKNFOTLKIIVILGLVLPVLMVICYSGILKTLRCRSEKGRHRAVLIF 238
QY 125 -----GGSYKCGLC-----QEFFGLNCCSSNRLDGHORV 154
Db 239 TIMIVYFLFWAPYINVLNLTQEFFGLNCCSSNRLDQAMQV 281

```

Search completed: October 21, 2003, 17:53:39
Job time : 65.9288 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 17:39:39 ; Search time 64.9288 Seconds
(without alignments)
383.806 Million cell updates/sec

Title: US-10-057-890A-31

Perfect score: 884

Sequence: i MKVSVAALSCMLVTALGSM.....GLNNCSSNRLDGHQVHAA 157

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_195Jun03.*

```
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	884	100.0	157	23	ABG32540 Human CCR5-based s
2	797	90.2	138	23	ABG32539 Human CCR5-based s
3	380	43.0	439	20	AA141280 Fusion protein con
4	379	42.9	352	18	AAW27407 Human CCR5. Homo
5	379	42.9	352	18	AAW27123 Human chemokine re
6	379	42.9	352	19	AAW23835 Human CC chemokine
7	379	42.9	352	20	AAW88232 HIV-1 co-receptor
8	379	42.9	352	22	ABW53142 Non-endogenous hum
9	379	42.9	352	22	AAG80111 Human CCR5 protein

10	379	42.9	352	22	AAB82948 Human HIV-1 co-rec
11	379	42.9	352	22	AAE07039 Human G-protein ch
12	379	42.9	352	22	AAE07048 Human G-protein ch
13	379	42.9	352	22	AAB83354 Human CCR5 protein
14	379	42.9	352	22	AAE04321 Human chemokine re
15	379	42.9	352	22	AAB46858 Human HDGMR10 prot
16	379	42.9	352	23	AAE25811 Human G-protein ch
17	379	42.9	352	23	ABG70597 Human G-protein ch
18	379	42.9	352	23	ABG92883 Human immunoglobul
19	379	42.9	352	23	ABB81054 G-protein chemokine
20	379	42.9	352	23	AAU97152 Human G-protein ch
21	379	42.9	352	23	ABB08343 Human chemokine (C
22	379	42.9	352	23	AAW52828 Human CC chemokine
23	379	42.9	352	23	AAW52829 Human CCR5 Gln 55
24	379	42.9	352	24	ABP97728 Amino acid sequenc
25	379	42.9	352	24	ABG75540 Human G-protein ch
26	379	42.9	352	24	ABP81933 Human C-C chemokine
27	374	42.3	371	19	AAW23834 Human CC chemokine
28	373	42.2	352	18	AAW07602 Human G-protein ch
29	373	42.2	352	21	AAW80128 Human G-protein ch
30	373	42.2	352	22	AAE07037 Human G-protein ch
31	373	42.2	352	22	AAE07046 Human G-protein ch
32	373	42.2	352	23	AAE25808 Human G-protein ch
33	373	42.2	352	23	ABG92880 Human G-protein ch
34	373	42.2	352	23	AAU97150 Human G-protein ch
35	363	41.1	352	18	AAW27125 Macaque chemokine
36	359	40.6	352	22	AAW79089 Amino acid sequenc
37	272	30.8	354	19	AAW54037 Mouse CC-CR5 prot
38	258	29.2	184	18	AAW27406 Inactive human CCR
39	258	29.2	215	18	AAW27408 HIV-1 co-receptor
40	258	29.2	215	20	AAW88238 Human chemokine re
41	183.5	20.8	332	18	AAW26766 HIV-1 co-receptor
42	182.5	20.6	100	20	AAW88231 HIV-1 co-receptor
43	182	20.6	32	22	ABG80086 Chemokine CCR5 ext
44	177	20.0	32	19	AAW43018 Synthetic peptide
45	177	20.0	32	19	AAW39912 Peptide representi

ALIGNMENTS

RESULT 1
ABG32540
ID ABG32540 standard; protein; 157 AA.
XX AC ABG32540;
XX DT 15-NOV-2002 (first entry)
XX XX Human CCR5-based scaffolded fusion protein #2.

XX Scaffolded protein; CCR5; HIV; human immunodeficiency virus infection;
XX ECD; extracellular domain; metal chelating motif; zinc finger protein;
XX integral membrane protein; soluble loop; intracellular domain; ICD;
XX gene therapy; immunogen; viral infection; human.

XX Homo sapiens.
OS Synthetic.

XX Key Location/Qualifiers
XX Peptide 1..19
XX Protein /label= Signal_peptide
XX FT 20..157
XX FT /label= Mature_scaffolded_protein

XX WO200260477-A1.

XX 08-AUG-2002.

XX 29-JAN-2002; 2002WO-US02377.

XX 31-JAN-2001; 2001US-265782P.

XX 31-JAN-2001; 2001US-265858P.

QY 140 NCSSNLDGCHORVHA 157
 DB 121 NCSSNLDGCHORVHA 138

RESULT 3
 ID AAY41280
 AC AAY41280; Protein; 439 AA.
 DT 11-FEB-2000 (first entry)
 DE Fusion protein containing human chemokine receptor CCR-5.
 KW Prostaglandin, PG; E2EP3 receptor; E2EP2 receptor; CCR-5; human; chemokine receptor; ss2 adrenergic receptor; small G-protein rho; renal outer medullary potassium ion channel protein; ion-channel protein; lambda phage repressor protein; G-protein coupled receptor; bacteria; biochemical; vaccine; immunohistochemical; orphan receptor; HIV.
 OS Synthetic.
 OS Homo sapiens.
 PN WO9953033-A1.
 PD 21-OCT-1999.
 PF 16-APR-1999; 99WO-US08214.
 PR 16-APR-1998; 98US-0081989.
 PA (UYVA-) UNIV VANDERBILT.
 PI Breyer RM, Ma L, Kennedy C;
 DR WPI; 1999-620416/53.
 DR N-PSDB; AAZ24738.
 KW New nucleic acid constructs for high level expression of eukaryotic proteins in bacteria, for producing e.g. chemokine receptor CCR-5 for preventing HIV infection
 PT Disclosure; Page 49-53; 81pp; English.
 PS
 CC The invention provides isolated nucleic acid sequences that encode rabbit prostaglandin (PG) E2EP3 receptor, human PG E2EP2 receptor, human chemokine receptor CCR-5, human ss2 adrenergic receptor, rat renal outer medullary potassium ion channel protein or human small G-protein rho, together with deduced protein sequences. Also provided is a method for the production of eukaryotic proteins by culturing bacteria transformed with vectors containing the above nucleic acid sequences or a nucleic acid (1) that comprises: (i) first sequence that encodes either a sequence comprising at least three positively charged amino acids, or a DNA-binding protein, or a lambda phage repressor protein, placed upstream of, and in frame with, (ii) a sequence encoding a protein. (i) are used for recombinant production of eukaryotic proteins, particularly membrane proteins, G-protein coupled receptors or ion-channel proteins, in bacteria. These proteins are useful for biochemical or structural studies; as therapeutic agents; in diagnostic and screening assays and as antigens for use in vaccines, and for raising antibodies that are useful as immunohistochemical markers, e.g. for orphan receptors or ion channels. Antibodies raised against the chemokine receptor CCR-5 can be used (when administered as antiserum or generated in vivo) to prevent entry of human immune deficiency virus (HIV) into cells.
 XX Sequence 439 AA;
 SQ

QY 1 MKVSVAALSCLMLVTALGSMYQVSPYDINYYTSEPCQKINVKQIAA----- 49
 DB 70 LKVSVEEF-----HMDYQVSPYDINYYTSEPCQKINVKQIAARLLPPLYSLVF 119
 QY 50 -----YKGLC-----AAQWDFGNTMC 67
 DB 120 IFGFVGNMLVILLINCKLKSMTDIYLLNLAISDLFFELLTVPFWAHYAAQWDFGNTMC 179
 QY 68 QHQRVGHGHHHSYKCG----- 83
 DB 180 --QLLTGLYFIGFFSGIFILLITIDRYLVAVHAFALKARTVTGVTSTVITWVAVFA 237
 QY 84 -----LCTRSQKGLHYTCSSHPYSOYQFKNQFOTLKI----- 117
 DB 238 SLPGIIFTRSQKGLHYTCSSHPYSOYQFKNQFOTLKIIVILGLVLPLLMVVCYSGILK 297
 QY 118 -----HORVHGG-----GGSYKGLC-----QEFGNLCSSSNRLDGHQ 152
 DB 298 TELRCRNEKKRHRAVLRTITMIVVFLFWAPYIVLLNTTQEFFGLNCCSSNRLDQAM 357
 QY 153 RV 154
 DB 358 QV 359

RESULT 4
 ID AAW27407 standard; Protein; 352 AA.
 AC AAW27407;
 DT 14-APR-1998 (first entry)
 DE Human CCR5.
 KW Human Cys-Cys chemokine receptor 5; CCR5;
 KW human immunodeficiency virus; type 1; type 2; HIV-1; HIV-2;
 KW diagnosis; treatment; prevention;
 KW inflammatory disease; rheumatoid arthritis; glomerulonephritis;
 KW asthma; idiopathic pulmonary fibrosis; psoriasis; viral infection;
 KW cancer; atherosclerosis; autoimmune disorder.
 OS Homo sapiens.
 PN WO9732019-A2.
 PD 04-SEP-1997.
 PF 28-FEB-1997; 97WO-BE00023.
 PR 06-AUG-1996; 96EP-0870102.
 PR 01-MAR-1996; 96EP-0870021.
 XX (EURO-) EUROSREEN SA.
 XX Libert F, Parmentier M, Samson M, Vassart G;
 WPI; 1997-479829/44.
 DR N-PSDB; AAT90117.
 XX Active and inactive forms of human CC chemokine receptor CCR-5 - useful to diagnose, prevent and/or treat inflammatory disorders, autoimmune disease and viral infection
 PS Claim 4; Fig 1b-c; 94pp; English.
 CC The present sequence is human CC (Cys-Cys) chemokine receptor 5 (CCR5), which is stimulated by MIP-1 alpha, MIP-1 beta or RANTES chemokines, but not by monocyte chemoattractant protein 1 (MCP-1), MCP-2, MCP-3, interleukin-8 (IL-8) or growth related gene product of alpha (GRO alpha) chemokines. Active CCR-5 is also a receptor of human immunodeficiency virus type 1 or type 2 (HIV-1 or HIV-2). CCR5 or its cDNA can be used to diagnose, treat and/or prevent

Query Match 43.0%; Score 380; DB 20; Length 439;

Best Local Similarity 34.1%; Pred. No. 4.5e-31;

Matches 103; Conservative 8; Mismatches 31; Indels 160; Gaps 8;

CC inflammatory diseases, e.g. rheumatoid arthritis,
 CCC glomerulonephritis, asthma, idiopathic pulmonary fibrosis and
 CCC psoriasis, viral infections, especially HIV-1 or HIV-2 infection,
 CC cancer, atherosclerosis and autoimmune disorders.

SQ Sequence 352 AA;

Query Match 42.9%; Score 379; DB 18; Length 352;
 Best Local Similarity 35.0%; Pred. No. 4.4e-31;
 Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;
 2Y 20 MDYQVSSPIYDINYYTSPCKINVKQIAA----- 49
 1 MDYQVSSPIYDINYYTSPCKINVKQIAARLLPLSLVIFGVGNMLVILLINCKR 60
 2Y 50 -----YKGLC-----AAAQWDFGNTMCQHORVGHGHHHSYKCG--- 83
 61 LKSMTDIYLLNLALISDLFFLLTPVFWAHYAAQWDFGNTMC--QLLTGLYPIGFSGIFF 118
 2Y 84 -----LCTRSQKEGLHYTC 97
 119 IILLTIDRYLAVVHAFVALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178
 2Y 98 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 124
 179 SSHFPYSQYQFWKNFQTLKIVILGLVPLLVVVICYSGLKTLRCRNEKRRHRAVLIF 238
 2Y 125 -----GGSYKGLC-----QEFGFLNCCSSNRLDGHORV 154
 239 TIMIVYFLWAPYINIVLLNTFQEFGLNCCSSNRLDQAMQV 281

RESULT 5

AAW27123
 D AAW27123 standard; Protein; 352 AA.

CC AAW27123;

T 14-DEC-1997 (first entry)

E Human chemokine receptor 88C.

X Chemokine receptor 88C; atherosclerosis; rheumatoid arthritis;
 X tumour; asthma; viral infection; AIDS; inflammation;
 W autoimmune disease; therapy; diagnosis; leukocyte trafficking;
 W G protein coupled receptor; ligand; modulator; antibody; human.

CS Homo sapiens.

Key Location/Qualifiers

T Domain 1..32
 T /label= Extracellular_domain
 T Domain 56..67
 T /label= Intracellular_domain
 T Domain 89..112
 T /label= Extracellular_domain
 T Domain 125..145
 T /label= Intracellular_domain
 T Domain 166..191
 T /label= Extracellular_domain
 T Domain 213..235
 T /label= Intracellular_domain
 T Domain 259..280
 T /label= Extracellular_domain
 T Domain 301..352
 T /label= Intracellular_domain

X W09722698-A2.

X 26-JUN-1997.

X 20-DEC-1996; 96WO-US20759.

PR 07-JUN-1996; 96US-0661393.
 PR 20-DEC-1995; 95US-0575967.
 XX (ICOS-) ICOS CORP.
 XX Gray PW, Raport CJ, Schweickart VL;
 XX WPI; 1997-341689/31.
 DR N-PSDB; AAT85161.

XX New nucleic acid encoding chemokine receptors 88-2B and 88C - used
 PT to modulate leukocyte trafficking, e.g. for treatment of
 PT inflammation, tumours, viral infections, autoimmune diseases, etc.
 XX Claim 16; Page 47-48; 65pp; English.

XX This polypeptide sequence comprises novel human chemokine receptor
 CC 88C, a G protein coupled receptor that is involved in leukocyte
 CC trafficking, its amino sequence was deduced from a cDNA clone
 CC (AAT85161) isolated from a macrophage library. It shows 62% identity
 CC to CCR1. Chemokine receptor 88-2B (see AAW27124) has also been
 CC identified. 88C and 88-2B receptors and their polypeptide fragments
 CC can be produced in transformed host cells. The receptors, peptides
 CC domains, and anti-receptor antibodies can be used to modulate
 CC receptor activities, particularly ligand and G protein binding, and
 CC are potentially useful in the treatment of
 CC atherosclerosis, rheumatoid arthritis, tumours, asthma, viral
 CC infection, AIDS, inflammatory conditions, pathological immune
 CC response, abnormal haematopoietic processes etc.

XX SQ Sequence 352 AA;

Query Match 42.9%; Score 379; DB 18; Length 352;

Best Local Similarity 35.0%; Pred. No. 4.4e-31;
 Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

QY 20 MDYQVSSPIYDINYYTSPCKINVKQIAA----- 49
 DB 1 MDYQVSSPIYDINYYTSPCKINVKQIAARLLPLSLVIFGVGNMLVILLINCKR 60
 QY 50 -----YKGLC-----AAAQWDFGNTMCQHORVGHGHHHSYKCG--- 83
 DB 61 LKSMTDIYLLNLALISDLFFLLTPVFWAHYAAQWDFGNTMC--QLLTGLYPIGFSGIFF 118
 QY 84 -----LCTRSQKEGLHYTC 97
 DB 119 IILLTIDRYLAVVHAFVALKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178
 QY 98 SSHFPYSQYQFWKNFQTLKI-----HQRVHGG----- 124
 DB 179 SSHFPYSQYQFWKNFQTLKIVILGLVPLLVVVICYSGLKTLRCRNEKRRHRAVLIF 238
 QY 125 -----GGSYKGLC-----QEFGFLNCCSSNRLDGHORV 154
 DB 239 TIMIVYFLWAPYINIVLLNTFQEFGLNCCSSNRLDQAMQV 281

RESULT 6

AAW23835

ID AAW23835 standard; Protein; 352 AA.

XX AAW23835;

XX 08-JUN-1998 (first entry)

DE Human CC chemokine receptor 5 (CCR5).

XX CC chemokine receptor 5; CCR5; G-protein coupled receptor;
 KW human immunodeficiency virus; HIV; CD4; AIDS; therapy;
 KW transgenic animal.

XX Homo sapiens.

XX FH Key Location/Qualifiers
 FT Domain 29..55
 FT /label= I
 FT Region 109..120
 FT Domain /note= "extracellular loop-1 (Claim 19)"
 FT /label= III
 FT Domain 104..126
 FT /note= "transmembrane domain"
 FT /label= IV
 FT Domain 143..171
 FT /note= "transmembrane domain"
 FT /label= V
 FT Region 187..210
 FT Domain /note= "extracellular loop-2 (Claim 19)"
 FT /label= VI
 FT Domain 238..258
 FT /note= "transmembrane domain"
 FT /label= VII
 FT Region 261..276
 FT Domain /note= "extracellular loop-3 (Claim 19)"
 FT /label= VIII
 FT /note= "transmembrane domain"
 FT
 FT WO9745543-A2.
 XX
 XX 04-DEC-1997.
 XX
 XX 28-MAY-1997; 97WO-US09586.
 XX
 XX 28-MAY-1996; 96US-0018508.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 XX Alkhatib G, Berger EA, Broder CC, Combadiere C;
 PI Feng Y, Kennedy PE, Murphy PM;
 XX
 XX WPI; 1998-032650/03.
 DR N-PSDB; AAT76920.
 XX
 XX CC Chemokine receptor 5 polypeptide - used to inhibit membrane
 PT fusion between HIV and a target cell
 XX
 XX Claim 68; Fig 1C; 70pp; English.
 XX
 CC This protein sequence comprises of a novel human macrophage-selective
 CC chemokine receptor that has been designated CCR5. The sequence
 CC was deduced from an isolated cDNA clone (see AAT76920). An Alal271eu
 CC variant (see W238340 of CCR5 was also identified. The susceptibility
 CC of human macrophages to HIV infection depends on cell surface
 CC expression of CD4 and CCR5. CCR5 is a member of the 7-transmembrane
 CC superfamily of G-protein coupled cell surface molecules. It plays
 CC an essential role in the membrane fusion step of infection by some
 CC HIV isolates. The establishment of stable, non-human cell lines
 CC and transgenic mammals having cells that coexpress human CD4 and
 CC CCR5 provides valuable tools for research of HIV infection.
 CC Antibodies that bind to CCR5, CCR5 variants, and CCR5-binding
 CC agents capable of blocking membrane fusion between HIV and target
 CC cells represent potential anti-HIV therapeutics for macrophage
 CC tropic strains of HIV.
 XX
 XX Sequence 352 AA;
 SQ
 Query Match 42.9%; Score 379; DB 19; Length 352;
 Best Local Similarity 35.0%; Pred. No. 4.4e-31;
 Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;
 20 MDYQVSSPIYDINYTSEPCQKINVKQIAA----- 49
 1 MDYQVSSPIYDINYTSEPCQKINVKQIAARLLPPLYSIVFTFGVGNMVLILLINCKR 60

QY 50 -----YKCGLC-----AAAQWDFGNTMCQHQHVHGHHSYKCG--- 83
 DB 61 LKSMTDIYLLNLATISDLFFLLTVPFMAHYAAQWDFGNTMC--QLLTGLYFIGFSGIFF 118
 QY 84 -----LCTRSQKGLHYTC 97
 DB 119 IILLIDRYLAVVHAVPALKARTVTFGVTVSVITWVAVFASLPGIIFTRSQKGLHYTC 178
 QY 98 SSHFPYSQYQFWKQFQTLKI-----HQRVHGG----- 124
 DB 179 SSHFPYSQYQFWKQFQTLKIIVLGVLPLLVVICYSGLKTLRCRNEKXKRAVRLIF 238
 QY 125 -----GGSYKGLC-----QEFGGLNCCSSNRLDGHRV 154
 DB 239 TIMIVYFLFWAPYNIIVLLNTFQEFFGLNCCSSNRLDQAMQV 281
 RESULT 7
 AAW88232
 ID AAW88232 standard; Protein; 352 AA.
 XX
 XX AAW88232;
 XX AC
 XX 15-MAR-1999 (first entry)
 XX
 XX HIV-1 co-receptor CCR5.
 XX
 KW HIV-1; CCR5; CCR5m303; co-receptor; infection; diagnosis; AIDS;
 KW gene therapy; human.
 XX
 XX Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 32..56
 FT /note= "transmembrane domain 1"
 FT Domain 67..87
 FT /note= "transmembrane domain 2"
 FT Misc-difference 101
 FT /note= "corresponds to TGT (Cys) in wild-type CCR5,
 FT TGA (Stop) in CCR5m303"
 FT Domain 103..124
 FT /note= "transmembrane domain 3"
 FT Domain 142..167
 FT /note= "transmembrane domain 4"
 FT Domain 200..223
 FT /note= "transmembrane domain 5"
 FT Domain 236..260
 FT /note= "transmembrane domain 6"
 FT Domain 275..301
 FT /note= "transmembrane domain 7"
 XX WO9854317-A1.
 XX
 XX 03-DEC-1998.
 XX
 XX 29-MAY-1998; 98WO-EP03437.
 XX
 XX 30-MAY-1997; 97US-0048057.
 XX
 XX (MOND-) FOND MONDIALE RECH & PREVENTION SIDA.
 XX
 XX Arenzana Siesdedos F, Beretta A, Braun J, Quillent C;
 PI WPI; 1999-059835/05.
 XX N-PSDB; AAV84126.
 DR
 DR WPI; 1999-059835/05.
 XX
 XX New CCR5 variant protein of the HIV-1 co-receptor - useful in
 PT developing resistance of CCR5-expressing cells to HIV-1 infection
 XX
 XX Disclosure; Page 34-35; 55pp; English.
 XX
 XX This is the amino acid sequence of wild-type human CCR5, which

CC serves as a co-receptor for infection by macrophage-tropic
CC (M-tropic) strains of HIV-1. The invention relates to the
CC identification of a CCR5 variant (see AAW89231), designated CCR5m303,
CC comprising the first two transmembrane domains of wild-type CCR5,
CC but lacking transmembrane domains 3-7. The presence of the
CC CCR5m303 variant with the wild type CCR5 allele shows a positive
CC correlation with resistance to infection with M-tropic HIV-1
CC strains, and may indicate slower progression of the disease. The
CC detection of CCR5 variants may be used to identify individuals at
CC lower risk of infection relative to the general population who, if
CC infected, may exhibit slower progression to AIDS. Probes and
CC primers (see AAW84127-36) are provided for use in diagnostic methods
CC for detecting the presence of such variants. A method is provided
CC for inhibiting HIV-1 infection of a cell expressing the CCR5
CC receptor. This involves introducing a nucleic acid encoding a CCR5
CC variant into the cell, thereby reducing the number of functional
CC CCR5 molecules present on the cell surface.

XX Q Sequence 352 AA;

Query Match 42.9%; Score 379; DB 20; Length 352;
Best Local Similarity 35.0%; Pred. No. 4.4e-31;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

20 MDYQVSSPIYDINVTSPCKINVKQIAA----- 49
|||||
1 MDYQVSSPIYDINVTSPCKINVKQIAARLLPPLYSLVFIFGVGNMLVILILINCKR 60
50 -----YKCGLC-----AAAQWDFGNTMCOHORVHGHHHSYKCG--- 83
|||||
61 LKSMTDIYLLNLAIISDLFFLLTVPFWAHYAAQWDFGNTMC--QLLTGLYFIFGFSGIF 118
84 -----LCTRSQKEGLHYTC 97
119 IILLIDRYLAVHNAVFAKARTVTFGVTVITWVAVFASLPGIIFTRSQKEGLHYTC 178
98 SSHFPYSQYQFQKNTQTLKI-----HORVHG----- 124
|||||
179 SSHFPYSQYQFQKNTQTLKIIVILGLVPLLVWVICYSGLIKTLRCRNEKKHRAKRLIF 238
125 -----GGSYKCGLC-----QEPFGLNCCSSNRLDGHORV 154
|||||
239 TIMIVYFLWAPYNIIVLLNTFQEPFGLNCCSSNRLDQAMQV 281

RESULT 8
ABB56342
ID ABB56342 standard; Protein; 352 AA.

XX ABB56342;

18-FEB-2002 (first entry)

XX Non-endogenous human GPCR protein, SEQ ID NO: 477.

XX Human; G protein-coupled receptor; GPCR; non-endogenous; mutant;
XX constitutively activated GPCR; agonist; disease.

XX Homo sapiens.
XX Synthetic.

XX WO200177172-A2.

XX 18-OCT-2001.

XX 05-APR-2001; 2001WO-US11098.

XX 07-APR-2000; 2000US-195747P.

XX (AREN-) ARENA PHARM INC.

XX Lehmann-Bruinsma K, Liaw CW, Lin I;

DR WPI: 2001-648759/74.
DR N-PSDB; ABI97978.

XX Identifying agonists of G protein-coupled receptors (GPCRs) for use in
PT disease treatment, comprises contacting candidate compounds with
PT versions of GPCRs -

XX Claim 1; Page 277-278; 394pp; English.

XX The invention relates to G protein-coupled receptors (GPCRs) for which
CC the endogenous ligand has been identified. Non-endogenous
CC constitutively activated versions of known GPCRs are used in the
CC invention for the direct identification of candidate compounds as
CC receptor agonists, inverse agonists or partial agonists. Such
CC agonists are useful as therapeutic agents for diseases or disorders
CC associated with GPCRs. The present sequence is a non-endogenous
CC version of a known human GPCR.

XX Q Sequence 352 AA;

Query Match 42.9%; Score 379; DB 22; Length 352;
Best Local Similarity 35.0%; Pred. No. 4.4e-31;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

20 MDYQVSSPIYDINVTSPCKINVKQIAA----- 49
|||||
1 MDYQVSSPIYDINVTSPCKINVKQIAARLLPPLYSLVFIFGVGNMLVILILINCKR 60
50 -----YKCGLC-----AAAQWDFGNTMCOHORVHGHHHSYKCG--- 83
|||||
61 LKSMTDIYLLNLAIISDLFFLLTVPFWAHYAAQWDFGNTMC--QLLTGLYFIFGFSGIF 118
84 -----LCTRSQKEGLHYTC 97
119 IILLIDRYLAVHNAVFAKARTVTFGVTVITWVAVFASLPGIIFTRSQKEGLHYTC 178
98 SSHFPYSQYQFQKNTQTLKI-----HORVHG----- 124
|||||
179 SSHFPYSQYQFQKNTQTLKIIVILGLVPLLVWVICYSGLIKTLRCRNEKKHRAKRLIF 238
125 -----GGSYKCGLC-----QEPFGLNCCSSNRLDGHORV 154
|||||
239 TIMIVYFLWAPYNIIVLLNTFQEPFGLNCCSSNRLDQAMQV 281

RESULT 9
AAG80111
ID AAG80111 standard; Protein; 352 AA.

XX AAG80111;

17-JAN-2002 (first entry)

XX Human CCR5 protein.

XX Chemokine; tumour diagnosis; colorectal; prostatic; organ rejection;
KW inflammation; autoimmune disease; metastasis; bronchial asthma; lupus;
KW chronic bowel inflammation; rheumatoid arthritis; cytostatic;
KW antiinflammatory; antiasthmatic; immunosuppressive; dermatological;
KW antirheumatic; antiarthritic.

XX Homo sapiens.

XX WO200172830-A2.

XX 04-OCT-2001.

XX 02-APR-2001; 2001WO-EP03708.

XX 31-MAR-2000; 2000DE-1016013.

XX (IPFP-) IPF PHARM GMBH.

XX (FORS/) FORSSMANN U.

XX PI Forssmann W, Adermann K, Heitland A, Spodeberg N;
 XX WPI; 2001-626256/72.
 XX
 XX Diagnostic agent containing two or more receptor-specific ligands,
 PT useful for detecting tumors, inflammation etc., also therapeutic use of
 PT ligand inhibitors -
 XX
 XX Disclosure; Page 10; 26pp; German.
 XX
 XX This invention describes a novel diagnostic agent (A) comprising at least
 CC two different ligands (I) for receptors (II) that are implicated in
 CC disease. (A) are used for the diagnosis of tumors (especially colorectal
 CC or prostatic), organ rejection, inflammation and autoimmune diseases.
 CC Also inhibitors of (I) are used therapeutically against tumors (and their
 CC metastases), inflammation (particularly bronchial asthma or chronic bowel
 CC inflammation), or autoimmune diseases (rheumatoid arthritis or lupus),
 CC where the (cardio)vascular, lymphatic, respiratory, nervous, digestive,
 CC endocrine, motor or urogenital systems or skin are affected, and bone
 CC marrow diseases. The products of the invention are chemokine derivatives
 CC which have cytostatic, antiinflammatory, antiaethmatic,
 CC immunosuppressive, dermatological, antirheumatic, antiarthritic.
 CC Chemokines act on specific tumor and inflammatory cells through a
 CC constellation of chemokine receptors (CR), which control migration and
 CC proliferation of these cells. AAG80045-AAG80128 represent human chemokine
 CC fragments used to illustrate the method of the invention.
 XX
 SQ Sequence 352 AA;
 Query Match 42.9%; Score 379; DB 22; Length 352;
 Best Local Similarity 35.0%; Pred. No. 4.4e-31;
 Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;
 QY 20 MDYQVSSPIYDINTYTSEPCQKINVKQIAA----- 49
 DB 1 MDYQVSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSLVFIQFVGNMLVILLINCKR 60
 QY 50 -----YKCGLC-----AAAQWDFGNTMCOHORVGHGHHHSYKCG--- 83
 DB 61 LKSMTDIYLLNLATSDLFLLTVFPWAHYAAQWDFGNTMC--QLLTGLYFIQFSGIFF 118
 QY 84 -----GGSYKCGLC-----QEPFGLNCCSSNRDLGDHORV 97
 DB 119 ILLTIDRYLAVHVAFAKARTVFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTC 178
 QY 98 SSHFPYSQYQFWKQNFQTLKI-----HORVHG----- 124
 DB 179 SSHFPYSQYQFWKQNFQTLKIIVILGLVPLVMVTCYSGILKTLRCRNEKRRHRAVLIF 238
 QY 125 -----GGSYKCGLC-----QEPFGLNCCSSNRDLGDHORV 154
 DB 239 TIMIVYFLWAPYNIIVLLNTFQEPFGLNCCSSNRDLQAMQV 281

RESULT 10
 AAB82948
 ID AAB82948 standard; Protein; 352 AA.
 XX
 XX AAB82948;
 XX
 XX 21-DEC-2001 (first entry)
 XX
 XX Human HIV-1 co-receptor CCR5.
 XX
 XX CCR5; chemokine; co-receptor; human immunodeficiency virus type 1;
 KW HIV-1; infection; therapy; vaccine; anti-HIV-1.
 XX
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FT Binding-site 2..18
 FT /note= "binds to HIV-1 gp120"

XX PN WO200164710-A2.
 XX PD 07-SEP-2001.
 XX
 XX 28-FEB-2001; 2001WO-US06699.
 XX
 XX 29-FEB-2000; 2000US-185667P.
 PR 19-MAY-2000; 2000US-205839P.
 PR 07-FEB-2001; 2001US-267231P.
 XX
 XX (PROG-) PROGENICS PHARM INC.
 PA (AARO-) AARON DIAMOND AIDS RES CENT.
 XX
 XX Dragic T, Olson WC;
 XX
 XX WPI; 2001-611273/70.
 DR N-PSDB; AAB26903.
 XX
 XX Novel compounds comprising specific amino acids within CCR5 (HIV 1
 PT co-receptor) amino terminal domain including negatively charged and two
 PT sulfated tyrosine residues is useful for treating HIV infection in
 PT humans
 XX
 XX Claim 1; Page 30; 163pp; English.
 XX
 XX The present sequence is that of human HIV-1 co-receptor CCR5.
 CC Amino acids 2-18 in the N-terminal region of CCR5 comprise an HIV-1
 CC gp120-binding site that determines the specificity of the
 CC interaction between CCR5 and HIV-1 gp120. Post-translational
 CC sulfation of the tyrosine residues in the CCR5 N-terminus is
 CC required for gp120 binding and may critically modulate the
 CC susceptibility of target cells to HIV-1 infection in vivo. The
 CC invention provides claimed sulfated peptides (see AAB82947) that
 CC are based on the CCR5 N-terminal region and which are effective
 CC for inhibiting HIV-1 binding to CCR5. These peptides are used in
 CC claimed methods of inhibiting HIV infection of CD4+ cells, of
 CC preventing CD4+ cells from becoming infected with HIV, of treating
 CC a subject whose CD4+ cells are infected with HIV, and of
 CC identifying an agent which inhibits binding of a CCR5 ligand to a
 CC CCR5 receptor. The methods may be carried out in a subject, a
 CC especially a human, infected (therapeutic method), not infected
 CC with HIV (prophylactic method), or in a subject who is not infected
 CC with, but has been exposed to, HIV.
 XX
 SQ Sequence 352 AA;
 Query Match 42.9%; Score 379; DB 22; Length 352;
 Best Local Similarity 35.0%; Pred. No. 4.4e-31;
 Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;
 QY 20 MDYQVSSPIYDINTYTSEPCQKINVKQIAA----- 49
 DB 1 MDYQVSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSLVFIQFVGNMLVILLINCKR 60
 QY 50 -----YKCGLC-----AAAQWDFGNTMCOHORVGHGHHHSYKCG--- 83
 DB 61 LKSMTDIYLLNLATSDLFLLTVFPWAHYAAQWDFGNTMC--QLLTGLYFIQFSGIFF 118
 QY 84 -----GGSYKCGLC-----QEPFGLNCCSSNRDLGDHORV 97
 DB 119 ILLTIDRYLAVHVAFAKARTVFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTC 178
 QY 98 SSHFPYSQYQFWKQNFQTLKI-----HORVHG----- 124
 DB 179 SSHFPYSQYQFWKQNFQTLKIIVILGLVPLVMVTCYSGILKTLRCRNEKRRHRAVLIF 238
 QY 125 -----GGSYKCGLC-----QEPFGLNCCSSNRDLGDHORV 154
 DB 239 TIMIVYFLWAPYNIIVLLNTFQEPFGLNCCSSNRDLQAMQV 281

RESULT 11

AAE07039	Matches	99; Conservative	7; Mismatches	27; Indels	150; Gaps	7;
ID AAE07039 standard; Protein; 352 AA.	QY	20 MDYQVSSPIYDINYTSEPCQKINVKQIAA	-----	-----	-----	49
AAE07039;	DB	1 MDYQVSSPIYDINYTSEPCQKINVKQIAARLLPPLSLVLFVFGVGNMLVILILINCKR	-----	-----	-----	60
DT 16-OCT-2001 (first entry)	QY	50 -----YKCGLC-----	-----	-----	-----	83
Human G-protein chemokine receptor (CCR5) HDGNR10 protein #2.	DB	61 LKSMTDIYLLNLAISDLFFLLTVFPFAHAAQAQDFGNTMC--QLLTGLYFIFGFSGIF	-----	-----	-----	118
Human; G-protein chemokine receptor; CCR5; HDGNR10; inflammation; HIV;	QY	84 -----	-----	-----	-----	97
human immunodeficiency virus; antimicrobial; vasodilator; vulnerary;	DB	119 ILLTIDRYLAVHVAFAKXARTVTFGVTSVITWVAVPAFLPGIIFTRSQEGLHYTC	-----	-----	-----	178
cytostatic; immunosuppressive; neutropic; neuroprotective; gene therapy;	QY	98 SHHPYSQYQFQNPQTLKI-----	-----	-----	-----	124
neurodegenerative disorder; Kaposi's sarcoma; autoimmune disease;	DB	179 SHHPYSQYQFQNPQTLKIIVLGLVPLLVVVICYSGLKTLRLCRNKKRRAVRLIF	-----	-----	-----	238
rheumatoid arthritis; cancer; breast; ovary; adrenal gland; bone marrow;	QY	125 -----GGSYKCGLC-----	-----	-----	-----	154
gastrointestinal tract; lung; liver; immune disorder; Addison's disease;	DB	239 TIMIVYFLFWAPYNIVLLNTFQEFFGLNCCSSNRLLDQAMQV	-----	-----	-----	281
haemolytic anaemia; autoimmune thyroiditis; diabetes mellitus; allergy;	QY	-----	-----	-----	-----	-----
multiple sclerosis; ulcerative colitis; Crohn's disease; wound healing;	DB	-----	-----	-----	-----	-----
cardiovascular disorder; myocardial ischaemia.	QY	-----	-----	-----	-----	-----
Homo sapiens.	DB	-----	-----	-----	-----	-----
WO200158915-A2.	QY	-----	-----	-----	-----	-----
16-AUG-2001.	DB	-----	-----	-----	-----	-----
09-FEB-2001; 2001WO-US04152.	QY	-----	-----	-----	-----	-----
09-FEB-2000; 2000US-0181258.	DB	-----	-----	-----	-----	-----
09-MAR-2000; 2000US-0187999.	QY	-----	-----	-----	-----	-----
22-SEP-2000; 2000US-0234336.	DB	-----	-----	-----	-----	-----
(HUMA-) HUMAN GENOME SCI INC.	QY	-----	-----	-----	-----	-----
Rosen CA, Roschke V, Li Y, Ruben SM;	DB	-----	-----	-----	-----	-----
WPI; 2001-488965/53.	QY	-----	-----	-----	-----	-----
N-PSDB; AAD13198.	DB	-----	-----	-----	-----	-----
Isolated nucleic acid encoding a human G-protein chemokine receptor	QY	-----	-----	-----	-----	-----
(CCR5) HDGNR10 polypeptide, useful for preventing or treating	DB	-----	-----	-----	-----	-----
autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative	QY	-----	-----	-----	-----	-----
disorders and neurodegenerative disorders -	DB	-----	-----	-----	-----	-----
Example 40; Page 486-487; 495pp; English.	QY	-----	-----	-----	-----	-----
The present sequence is human G-protein chemokine receptor (CCR5) HDGNR10	DB	-----	-----	-----	-----	-----
protein. CCR5 HDGNR10 antibodies are useful for treating, preventing	QY	-----	-----	-----	-----	-----
or ameliorating a disease or disorder associated with inflammation,	DB	-----	-----	-----	-----	-----
defective or aberrant chemotaxis of immune cells, HIV infection (such as	QY	-----	-----	-----	-----	-----
Pneumocystis carinii pneumonia or Kaposi's sarcoma) or defective or	DB	-----	-----	-----	-----	-----
aberrant T-cell antigen presenting cell interaction. The disease or	QY	-----	-----	-----	-----	-----
disorder may also be an infectious disease (e.g. a viral infection such	DB	-----	-----	-----	-----	-----
as an early stage HIV infection, a cytomegalovirus infection, or a	QY	-----	-----	-----	-----	-----
poxvirus infection), an autoimmune disease (e.g. rheumatoid arthritis) or	DB	-----	-----	-----	-----	-----
a neurodegenerative disorder. The disease or disorder may be associated	QY	-----	-----	-----	-----	-----
with aberrant CCR5 expression, lack of CCR5 function, aberrant CCR5	DB	-----	-----	-----	-----	-----
ligand expression, or lack of CCR5 ligand function. CCR5 HDGNR10 protein	QY	-----	-----	-----	-----	-----
is used as a food additive or preservative to increase or decrease	DB	-----	-----	-----	-----	-----
storage capabilities. CCR5 HDGNR10 DNA are useful for chromosome	QY	-----	-----	-----	-----	-----
identification and in gene therapy. CCR5 HDGNR10 DNA, protein,	DB	-----	-----	-----	-----	-----
antibodies, agonists and antagonists are also useful in the	QY	-----	-----	-----	-----	-----
diagnosis, treatment and prevention of cancer (breast, ovary, adrenal	DB	-----	-----	-----	-----	-----
gland, bone, bone marrow, gastrointestinal tract, liver, lung,	QY	-----	-----	-----	-----	-----
urogenital); immune disorders (Addison's disease, allergies, autoimmune	DB	-----	-----	-----	-----	-----
haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's	QY	-----	-----	-----	-----	-----
disease, multiple sclerosis, rheumatoid arthritis and ulcerative	DB	-----	-----	-----	-----	-----
colitis); cardiovascular disorders (myocardial ischaemias) and	QY	-----	-----	-----	-----	-----
wound healing.	DB	-----	-----	-----	-----	-----
Sequence 352 AA;	QY	-----	-----	-----	-----	-----
Query Match 42.9%; Score 379; DB 22; Length 352;	DB	-----	-----	-----	-----	-----
Best Local Similarity 35.0%; Pred. No. 4.4e-31;	QY	-----	-----	-----	-----	-----

CC cells, HIV infection (such as Pneumocystis carinii pneumonia or Kaposi's
 CC sarcoma) or defective or aberrant T-cell antigen presenting cell
 CC interaction. The disease or disorder may also be an infectious disease
 CC (e.g. a viral infection such as an early stage HIV infection, a
 CC cytomegalovirus infection, or a poxvirus infection), an autoimmune
 CC disease (e.g. rheumatoid arthritis) or a neurodegenerative disorder. The
 CC disease or disorder may be associated with aberrant CCR5 expression, lack
 CC of CCR5 function, aberrant CCR5 ligand expression, or lack of CCR5 ligand
 CC function. CCR5 HDGNR10 protein is used as a food additive or preservative
 CC to increase or decrease storage capabilities. CCR5 HDGNR10 DNA are useful
 CC for chromosome identification and in gene therapy. CCR5 HDGNR10 DNA,
 CC protein, antibodies, agonists and antagonists are also useful in the
 CC diagnosis, treatment and prevention of cancer (breast, ovary, adrenal
 CC gland, bone, bone marrow, gastrointestinal tract, liver, lung,
 CC urogenital); immune disorders (Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis); cardiovascular disorders (myocardial ischaemias) and
 CC wound healing. The present sequence is human CCR5 HDGNR10 protein.
 XX Sequence 352 AA;

Query Match 42.9%; Score 379; DB 22; Length 352;
 Best Local Similarity 35.0%; Pred. No. 4.4e-31;
 Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;
 QY 20 MDYQVSSPIYDINTYTSEPCOKINVKQIAA----- 49
 DB 1 MDYQVSSPIYDINTYTSEPCOKINVKQIAARLLPPLYSLVFIFGVGNMLVILLINCKR 60
 QY 50 -----YKCGLC-----AAAQWDFGNTWCQHORVGHGHHHSYKCG--- 83
 DB 61 LKSMTDIYLLNLAISSLDFLLTVFPWAHYAAQWDFGNTWC--QLLTGLYFIFGFSGIF 118
 QY 84 ----- 178
 DB 119 ILLTIDRYLAVHVAFAKARTVTFGVTSVITWVAVFASLPGLIFTRSQEGLHYTC 178
 QY 98 SSHFPYSQYQFWKQFQTLKI-----HQRVHGG----- 124
 DB 179 SSHFPYSQYQFWKQFQTLKIVILGLVPLLVWVVCYSGLILKTLRCNKKRRAVLIF 238
 QY 125 -----GGSYKCGLC-----QEFFGLNCCSSNRDLGHQV 154
 DB 239 TIMIVYFLFWAPYNTVLLNTFQBFPGNLNCCSSNRDLQAMQV 281

RESULT 13
 AAB83354
 ID AAB83354 standard; Protein; 352 AA.
 XX
 AC AAB83354;

DT 09-OCT-2001 (first entry)

XX Human CCR5 protein sequence.

XX Chemotactic chemokine receptor 5; gp120; CD4; therapy; HIV; CCR5;
 KW human immunodeficiency virus; anti-inflammatory disease; human.

XX Homo sapiens.

XX EP111858-A2.

XX 25-JUL-2001.

XX 03-JAN-2001; 2001EP-0300020.

XX 12-JAN-2000; 2000GB-0000659.

PR 12-JAN-2000; 2000GB-0000661.

PR 12-JAN-2000; 2000GB-0000663.

XX (PFIZ) PFIZER LTD.

PA (PFIZ) PFIZER INC.
 XX
 PI Dobbs S, Perros M, Rickett GA;
 XX
 DR WPI; 2001-477088/52.
 DR N-PSDB; AAF87099.
 XX
 PT Determining if an agent can modulate CCR5-gp120 interaction, comprises
 PT incubating the agent with CCR5 and gp120 and determining if the agent
 PT modulates the interaction -
 XX
 PS Claim 1; Page 110; 113pp; English.
 CC This sequence represents the human CCR5 protein sequence.
 CC The invention relates to a method for determining whether an agent is
 CC capable of modulating the interaction of chemotactic chemokine receptor 5
 CC (CCR5) with gp120, comprising incubating the agent with CCR5 and gp120
 CC and determining whether the agent modulates the interaction, where gp120
 CC is associated with CD4, and where the interaction is a low affinity
 CC binding. The method is used to identify an agent capable of modulating
 CC the interaction of CCR5 with gp120. An agent identified by the method
 CC is used to prepare a pharmaceutical composition for the treatment of a
 CC disease or condition associated with CCR5 and gp120 interaction, to treat
 CC a subject with a disease or condition associated with CCR5 and gp120
 CC interaction, and for preparing a pharmaceutical for treating human
 CC immunodeficiency virus (HIV). It can also be used to treat anti-inflammatory
 CC diseases. The method is commercially useful, amenable to high throughput
 CC screening, and detects interaction of gp120 with cells expressing only
 CC CCR5.
 XX

SQ Sequence 352 AA;

Query Match 42.9%; Score 379; DB 22; Length 352;

Best Local Similarity 35.0%; Pred. No. 4.4e-31;

Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

QY 20 MDYQVSSPIYDINTYTSEPCOKINVKQIAA----- 49

DB 1 MDYQVSSPIYDINTYTSEPCOKINVKQIAARLLPPLYSLVFIFGVGNMLVILLINCKR 60

QY 50 -----YKCGLC-----AAAQWDFGNTWCQHORVGHGHHHSYKCG--- 83

DB 61 LKSMTDIYLLNLAISSLDFLLTVFPWAHYAAQWDFGNTWC--QLLTGLYFIFGFSGIF 118

QY 84 ----- 178

DB 119 ILLTIDRYLAVHVAFAKARTVTFGVTSVITWVAVFASLPGLIFTRSQEGLHYTC 178

QY 98 SSHFPYSQYQFWKQFQTLKI-----HQRVHGG----- 124

DB 179 SSHFPYSQYQFWKQFQTLKIVILGLVPLLVWVVCYSGLILKTLRCNKKRRAVLIF 238

QY 125 -----GGSYKCGLC-----QEFFGLNCCSSNRDLGHQV 154

DB 239 TIMIVYFLFWAPYNTVLLNTFQBFPGNLNCCSSNRDLQAMQV 281

RESULT 14

ID AAE04321 standard; Protein; 352 AA.

XX AAE04321;

XX 04-SEP-2001 (first entry)

XX Human chemokine receptor (CCR), CC-CCR-5 related protein #2.

XX Human; transformed mammalian cell; CD4; reporter gene; translocation;
 KW human immuno deficiency virus; HIV; long terminal repeat; LTR; therapy;
 KW chemokine receptor; CCR; cellular dysfunction; HIV infection; cofactor;
 KW CC-CCR-5; envelope glycoprotein; anti-HIV.

XX Homo sapiens.

XX US6258527-B1.
 DT 10-JUL-2001.
 DT 21-MAY-1997; 97US-0861105.
 XX 20-MAY-1996; 96US-0017157.
 XX 19-JUN-1996; 96US-0020043.
 XX 19-MAY-1997; 97US-0858660.
 XX (AARO-) AARON DIAMOND AIDS RES CENT.
 XX (UTNY) UNIV NEW YORK STATE.
 XX Littman DR, Deng H, Ellmeier W, Landau NR, Liu R;
 I WPI; 2001-417127/44.
 R N-PSDB; AAD08577.
 R Transformed mammalian cell (I) that contains a CD4 gene, reporter gene
 T and HIV LTR for identification of drugs and antibodies for treatment of
 T HIV -
 X Disclosure; Column 47-50; 37pp; English.
 X The present invention relates to a transformed mammalian cell that
 X contains a gene encoding CD4, a construct encoding a reporter gene under
 X the regulation of an human immuno deficiency virus (HIV) long terminal
 X repeat (LTR) and that has been transduced with a vector encoding a
 X human chemokine receptor (CCR) where the CD4 and the CCR are present on
 X the cell surface of transformed mammalian cell. The invention is useful
 X for identifying drugs or antibodies that interfere with the
 X translocation of HIV into transformed mammalian cell or for identifying
 X a human chemokine receptor that facilitates the infection of a
 X particular HIV strain into the transformed mammalian cell. Compounds
 X identified can be used to treat cellular dysfunction and to prevent or
 X combat HIV infection. The present sequence is a human chemokine receptor
 X (CCR), CC-CR-5 related protein. CC-CR-5 is the principal cofactor for
 X entry mediated by the envelope glycoproteins of primary macrophage-tropic
 X strains of HIV-1.
 X Q Sequence 352 AA;
 Query Match 42.9%; Score 379; DB 22; Length 352;
 Best Local Similarity 35.0%; Pred. No. 4.4e-31;
 Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;
 20 MDYQVSSPIYDINYYTSRPOKINVKQIAA----- 49
 1 MDYQVSSPIYDINYYTSRPOKINVKQIAARLLPPLYSLVIFGFGNMVLILINCKR 60
 50 -----YKGLC-----AAQMDFGNTMCOHQRVHGHHSYKCG--- 83
 61 LKSMTDIYLLNLAISDLPPLLPVFPWAHYAAQWDFGNTMC--QLLTGLYIFGFSGIF 118
 84 -----LCTRSQKGLHYTC 97
 119 IILLTIDRYLAVVHVAFAKARTVFGVVTSVITVWVAFSLPGIIFTRSQKGLHYTC 178
 98 SSHFPYSQYQFNKQFOTLKI-----HQRVHGG----- 124
 179 SSHFPYSQYQFNKQFOTLKIIVLGLVPLLVVVICYSGLKTLRCRNEKXRAVRLIF 238
 125 -----GGSYKGLC-----QEFFGLNCCSSNRLDGHQRV 154
 239 TIMIVVFLFPWAPYNIIVLLNTFQEFFGLNCCSSNRLDQAMQV 281
 RESULT 15
 D AAB46858
 X AAB46858 standard; Protein; 352 AA.
 X AAB46858;

XX DT 16-AUG-2001 (updated)
 DT 02-AUG-2001 (updated)
 DT 04-MAY-2001 (first entry)
 XX Human HDGNR10 protein.
 XX HDGNR10; human; G-protein chemokine receptor; antiinflammatory;
 KW immunomodulatory; anticoagulant; antiallergic; immunosuppressive;
 KW cytostatic; antiparasitic; antiporiatic; antirheumatic; antiarthritic;
 KW vasotropic; gene therapy; haematopoiesis; wound healing; coagulation;
 KW angiogenesis; solid tumour; infection; leukemia; growth factor activity;
 KW T-cell mediated autoimmune disease; psoriasis; allergy; atherogenesis;
 KW anaphylaxis; malignancy; inflammation; histamine; IGE; silicosis; shock;
 KW immunoglobulin E-mediated allergic reaction; rheumatoid arthritis;
 KW prostaglandin-independent fever; bone marrow failure; sarcoidosis;
 KW hyper-eosinophilic syndrome; vulneryary.
 XX Homo sapiens.
 OS US2001000241-A1.
 PN XX
 XX 12-APR-2001.
 PD XX
 XX 29-NOV-2000; 2000US-0725285.
 PF XX
 XX 06-JUN-1995; 95US-0466343.
 PR XX
 XX 18-NOV-1998; 98US-0195662.
 PR XX
 XX 25-JUN-1999; 99US-0339912.
 XX (LIYY/) LI Y.
 PA (RUBE/) RUBEN S M.
 PA Li Y, Ruben SM;
 XX WPI; 2001-226317/23.
 DR N-PSDB; AAF26390.
 XX
 FT New human G-protein chemokine receptor polypeptides and
 FT polynucleotides, useful for identifying (ant)agonists to the G-protein
 PT chemokine receptor -
 PT
 XX Claim 1a; Page 15; 22pp; English.
 PS This invention describes a novel receptor polypeptide (I) selected from
 XX (i) a fully defined 329 amino acid sequence (ii) fully disclosed in the
 CC specification; and (iii) a polypeptide encoded by the cDNA contained in a
 CC plasmid, and fragments, analogs and derivatives of the polypeptide. The
 CC products of the invention have antiinflammatory, immunomodulatory,
 CC anticoagulant, antiallergic, immunosuppressive, vulneryary, cytostatic,
 CC antiparasitic, antiporiatic, antirheumatic, antiarthritic and vasotropic
 CC activity and can be used for gene therapy. The G-protein chemokine
 CC receptors, HDGNR10, (I) are useful for screening for compounds which
 CC activate or inhibit activation of (I). The products of the invention can
 CC also be used for stimulating haematopoiesis, wound healing, coagulation,
 CC angiogenesis, treating solid tumours, chronic infections, leukemia,
 CC T-cell mediated autoimmune diseases, parasitic infections, psoriasis, and
 CC stimulating growth factor activity. HDGNR10 is useful for treating
 CC allergy, atherogenesis, anaphylaxis, malignancy, chronic and acute
 CC inflammation, histamine and immunoglobulin E (IgE)-mediated allergic
 CC reactions, prostaglandin-independent fever, bone marrow failure,
 CC silicosis, sarcoidosis, rheumatoid arthritis, shock and
 CC hyper-eosinophilic syndrome.
 CC (N.B. This record was resubmitted to correct errors in the keyword
 CC formatting).
 XX Sequence 352 AA;
 SQ Query Match 42.9%; Score 379; DB 22; Length 352;
 Best Local Similarity 35.0%; Pred. No. 4.4e-31;
 Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;
 20 MDYQVSSPIYDINYYTSRPOKINVKQIAA----- 49

```

|||||
Db 1 MDYQVSSPIYDINYTSBPCQKINVKQIAARLLPPLYSLVIFGFVGNMLVILINCKR 60
QY 50 -----YKCGLC-----AAQWDFGNTMCQHORVHGHHHSYKCG--- 83
Db 61 LKSMTDIYLLNLAISDLFFLLTVFFWAHYAAQWDFGNTMC--QLLTGLYFIFGFSGIF 118
QY 84 -----LCTRSQKEGLHYTC 97
Db 119 ILLTIDRYLAVHVAHFALKARTVTFGVVTSVITWVAVFASLPFGIIFTRSQKEGLHYTC 178
QY 98 SSHFPYSQYQFWKNFQTLKI-----HORVHG----- 124
Db 179 SSHFPYSQYQFWKNFQTLKIVILGLVPLLVMTVCYSGILKTLRCRNEKRRHRAVRLIF 238
QY 125 -----GGSYKGLC-----QEFFGLNCCSSNRLDGHRV 154
Db 239 TIMIVYFLFWAPYNIIVLLNTFQEFFGLNCCSSNRLDQAMQV 281
```

Search completed: October 21, 2003, 17:50:55
Job time : 65.9288 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

QM protein - protein search, using sw model

Run on: October 21, 2003, 17:53:49 ; Search time 47.9983 Seconds
(without alignments)
548.899 Million cell updates/sec

Title: US-10-057-890A-31

Perfect score: 884

Sequence: 1 MKVSVAALSCMLVLTALGSM.....GLNCCSSNRLDGHQVRHAA 157

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 629382 seqs, 167460630 residues

Total number of hits satisfying chosen parameters: 629382

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubaa/PCT PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubaa/PCT PUB PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubaa/US06 NEW PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubaa/US06 PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubaa/US07 NEW PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubaa/PCTUS PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubaa/US08 NEW PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubaa/US08 PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubaa/US09A PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubaa/US09B PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubaa/US09C PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubaa/US09 NEW PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubaa/US10A PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubaa/US10B PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubaa/US10C PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubaa/US10 NEW PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubaa/US60 NEW PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubaa/US60 PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	884	100.0	157	15	US-10-057-890A-31
2	797	90.2	138	15	US-10-057-890A-10
3	379	42.9	352	9	US-09-725-285-2
4	379	42.9	352	9	US-09-759-841-2
5	379	42.9	352	9	US-09-779-879A-22
6	379	42.9	352	9	US-09-779-880A-22
7	379	42.9	352	9	US-09-813-653-15
8	379	42.9	352	9	US-09-813-653-17
9	379	42.9	352	9	US-09-796-202-1
10	379	42.9	352	9	US-09-195-662A-2
11	379	42.9	352	9	US-09-339-912A-2
12	379	42.9	352	10	US-09-938-719-5
13	379	42.9	352	10	US-09-939-226-5
14	379	42.9	352	10	US-09-938-703-5
15	379	42.9	352	10	US-09-502-783A-2

16	379	42.9	352	11	US-09-734-221A-14	Sequence 14, Appl
17	379	42.9	352	12	US-10-323-314-1	Sequence 1, Appl
18	379	42.9	352	12	US-10-072-301-1	Sequence 1, Appl
19	379	42.9	352	12	US-10-071-866-1	Sequence 1, Appl
20	379	42.9	352	12	US-10-135-833-22	Sequence 22, Appl
21	379	42.9	352	12	US-10-239-423-67	Sequence 67, Appl
22	379	42.9	352	12	US-10-439-845-4	Sequence 4, Appl
23	379	42.9	352	14	US-10-106-623-2	Sequence 2, Appl
24	379	42.9	352	15	US-10-232-686-2	Sequence 2, Appl
25	379	42.9	352	15	US-10-086-814-1	Sequence 1, Appl
26	379	42.9	352	15	US-10-067-800-22	Sequence 22, Appl
27	379	42.9	352	15	US-10-290-058A-6	Sequence 6, Appl
28	379	42.9	352	15	US-10-225-567A-352	Sequence 352, Appl
29	374	42.9	352	12	US-10-164-649-52	Sequence 52, Appl
30	374	42.9	352	12	US-10-439-845-2	Sequence 2, Appl
31	373	42.2	352	9	US-09-779-879A-2	Sequence 2, Appl
32	373	42.2	352	9	US-09-779-880A-2	Sequence 2, Appl
33	373	42.2	352	12	US-10-135-839-2	Sequence 2, Appl
34	373	42.2	352	15	US-10-067-800-2	Sequence 2, Appl
35	363	41.1	352	14	US-10-106-623-20	Sequence 20, Appl
36	258	29.2	184	10	US-09-938-719-4	Sequence 4, Appl
37	258	29.2	184	10	US-09-939-226-4	Sequence 4, Appl
38	258	29.2	184	10	US-09-938-703-4	Sequence 4, Appl
39	258	29.2	215	10	US-09-938-719-6	Sequence 6, Appl
40	258	29.2	215	10	US-09-939-226-6	Sequence 6, Appl
41	258	29.2	215	10	US-09-938-703-6	Sequence 6, Appl
42	183.5	20.6	332	12	US-10-095-876A-2	Sequence 2, Appl
43	182	20.6	32	12	US-10-072-301-9	Sequence 9, Appl
44	182	20.6	32	12	US-10-071-866-9	Sequence 9, Appl
45	182	20.6	32	12	US-10-239-423-42	Sequence 42, Appl

ALIGNMENTS

RESULT 1

US-10-057-890A-31
; Sequence 31, Application US/10057890A
; Publication No. US20030044901A1
; GENERAL INFORMATION:
; APPLICANT: Coleman, Timothy
; APPLICANT: Mansfield, Brian
; TITLE OF INVENTION: Scaffold Fusion Polypeptides, Composition for Making the Same,
; FILE REFERENCE: PF537
; CURRENT APPLICATION NUMBER: US/10/057,890A
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: 60/265,782
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,858
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 31
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-890A-31

Query Match	100.0%	Score 884	DB 15	Length 157
Best Local Similarity	100.0%	Pred. No. 6.8e-80		
Matches 157	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	MKVSVAAALSCMLVLTALGSM	YVSSPIYDINYYTSEPCQKINVKQIAAYKGLCAAQW	60
DB	1	MKVSVAAALSCMLVLTALGSM	YVSSPIYDINYYTSEPCQKINVKQIAAYKGLCAAQW	60
QY	61	DFGNMTCQHORVGHGHHHSYKCGCLTRSOKEGLHYTCSHFPPYSOYQFWKNFQTLKIHOR		120
DB	61	DFGNMTCQHORVGHGHHHSYKCGCLTRSOKEGLHYTCSHFPPYSOYQFWKNFQTLKIHOR		120
QY	121	VHGGGGSYKGLCQCEFFGLNCCSSNRLDGHQVRHAA		157
DB	121	VHGGGGSYKGLCQCEFFGLNCCSSNRLDGHQVRHAA		157

RESULT 2

US-10-057-890A-10
Sequence 10, Application US/10057890A
Publication No. US20030044901A1
GENERAL INFORMATION:
APPLICANT: Coleman, Timothy
APPLICANT: Mansfield, Brian
TITLE OF INVENTION: Scaffold Fusion Polypeptides, Composition for Making the Same, and Use Thereof
TITLE OF INVENTION: of Using the Same.
FILE REFERENCE: PF537
CURRENT APPLICATION NUMBER: US/10/057.890A
CURRENT FILING DATE: 2002-01-29
PRIOR APPLICATION NUMBER: 60/265,782
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,858
PRIOR FILING DATE: 2001-01-31
NUMBER OF SEQ ID NOS: 32
SEQ ID NO 10
LENGTH: 138
TYPE: PRT
ORGANISM: Homo sapiens

US-10-057-890A-10

Query Match 90.2%; Score 797; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.4e-71;
Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
20 MDYQVSSPIYDINYTSEPCQKINVKQIAA-----AAQNDFGNTMCOHORVHGHHS 79
1 MDYQVSSPIYDINYTSEPCQKINVKQIAAARLLPPLYSLVFIFGVGNMVLILINCKR 60
80 YKCGICTRSQKGLHYTCSSHPFYSQYQFKNFOTLKIHQVHGGGSGYKGLCOEFFGL 139
61 YKCGICTRSQKGLHYTCSSHPFYSQYQFKNFOTLKIHQVHGGGSGYKGLCOEFFGL 120
140 NNCSSNRLDGHQVHAA 157
121 NNCSSNRLDGHQVHAA 138

RESULT 3

US-09-725-285-2
Sequence 2, Application US/09725285
Patent No. US2001000241A1
GENERAL INFORMATION:
APPLICANT: Li, Yi
APPLICANT: Ruben, Steven, M.
TITLE OF INVENTION: Antibodies to Human G-Protein Chemokine Receptor HDGFR10
TITLE OF INVENTION: (CCR5 Receptor)
FILE REFERENCE: 1488.1150003
CURRENT APPLICATION NUMBER: US/09/725.285
CURRENT FILING DATE: 2000-11-29
PRIOR APPLICATION NUMBER: 09/339,912
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: 09/195,662
PRIOR FILING DATE: 1998-11-18
PRIOR APPLICATION NUMBER: 08/466,343
PRIOR FILING DATE: 1995-06-06
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 352
TYPE: PRT
ORGANISM: Artificial Sequence: Genomic
FEATURE:
OTHER INFORMATION: Deduced Amino Acid Sequence

US-09-725-285-2

Query Match 42.9%; Score 379; DB 9; Length 352;
Best Local Similarity 35.0%; Pred. No. 1.5e-29;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

QY 20 MDYQVSSPIYDINYTSEPCQKINVKQIAA----- 49
DB 1 MDYQVSSPIYDINYTSEPCQKINVKQIAAARLLPPLYSLVFIFGVGNMVLILINCKR 60
QY 50 -----YKCGLC-----AAQNDFGNTMCOHORVHGHHSYKCG--- 83
DB 61 LKSMTDIYLLNLAIISDLFFLLTPVPFMAHYAAQWDFGNTMC--QLLTGLYFGFFSGIFF 118
QY 84 -----LCTRSQKGLHYTC 97
DB 119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVAVFASLPGLIFTRSQKGLHYTC 178
QY 98 SSHFPYSQYQFKNFOTLKI-----HORVHG----- 124
DB 179 SSHFPYSQYQFKNFOTLKIIVILGLVPLLVVICYSGLIKTLRCRNEKKHRAVRLIF 238
QY 125 -----GGSYKCGLC-----QEFFGLNCCSSNRLDGHQV 154
DB 239 TIMIVYFLFWAPYINIVLLNTFQEFFGLNCCSSNRLDQAMQV 281

RESULT 4

US-09-759-841-2
Sequence 2, Application US/C9759841
Patent No. US20010039026A1
GENERAL INFORMATION:
APPLICANT: Rickett, Graham A
APPLICANT: Dobbs, Susan
APPLICANT: Perros, Manoussos
TITLE OF INVENTION: Assay Method
FILE REFERENCE: PC10348APME
CURRENT APPLICATION NUMBER: US/09/759.841
CURRENT FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: GB 0000661.9
PRIOR FILING DATE: 2000-01-12
PRIOR APPLICATION NUMBER: GB 0000663.5
PRIOR FILING DATE: 2000-01-12
PRIOR APPLICATION NUMBER: GB 0000659.3
PRIOR FILING DATE: 2000-01-12
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 352
TYPE: PRT
ORGANISM: Homo sapiens
US-09-759-841-2

Query Match 42.9%; Score 379; DB 9; Length 352;
Best Local Similarity 35.0%; Pred. No. 1.5e-29;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

QY 20 MDYQVSSPIYDINYTSEPCQKINVKQIAA----- 49
DB 1 MDYQVSSPIYDINYTSEPCQKINVKQIAAARLLPPLYSLVFIFGVGNMVLILINCKR 60
QY 50 -----YKCGLC-----AAQNDFGNTMCOHORVHGHHSYKCG--- 83
DB 61 LKSMTDIYLLNLAIISDLFFLLTPVPFMAHYAAQWDFGNTMC--QLLTGLYFGFFSGIFF 118
QY 84 -----LCTRSQKGLHYTC 97
DB 119 IILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVAVFASLPGLIFTRSQKGLHYTC 178
QY 98 SSHFPYSQYQFKNFOTLKI-----HORVHG----- 124
DB 179 SSHFPYSQYQFKNFOTLKIIVILGLVPLLVVICYSGLIKTLRCRNEKKHRAVRLIF 238
QY 125 -----GGSYKCGLC-----QEFFGLNCCSSNRLDGHQV 154
DB 239 TIMIVYFLFWAPYINIVLLNTFQEFFGLNCCSSNRLDQAMQV 281

RESULT 5

US-09-779-879A-22
; Sequence 22, Application US/09779879A
; Patent No. US20020048786A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Roschke, Viktor
; APPLICANT: Li, Yi
; APPLICANT: Ruben, Steven, M.
; TITLE OF INVENTION: Human G-protein Chemokine Receptor (CCR5) HDGMR10
; FILE REFERENCE: 1488.115000A
; CURRENT APPLICATION NUMBER: US/09/779.879A
; CURRENT FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/181,258
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: US 60/187,999
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: US 60/234,336
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-779-879A-22

Query Match 42.9%; Score 379; DB 9; Length 352;
Best Local Similarity 35.0%; Pred. No. 1.5e-29;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;
QY 20 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 49
DB 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIQFVGNMLVILLINCKR 60
QY 50 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 83
DB 61 LKSMTDIYLLNLAISDLFFLLTVPFWAHAAQWDFGNTMC--QLLTGLYFIQFSGIFF 118
QY 84 ----- 124
DB 119 ILLTIDRYLAVHVAFAKARTVTFGVTSVITWVAVFASLPGLIFTRSQEGLHYTC 178
QY 98 SSHFPYSQYQFQKQFQTLKI-----HORVHGG----- 124
DB 179 SSHFPYSQYQFQKQFQTLKIIVILGLVPLVWVICYSGLKTLRCRNEKKRHRAVRLIF 238
QY 125 -----GSKYKGLC-----QFFGLNCCSSNRDLGQHV 154
DB 239 TIMIVYFLWAPYNIIVLLNTFQFFGLNCCSSNRDLQAMQV 281

RESULT 6

US-09-779-880A-22
; Sequence 22, Application US/09779880A
; Patent No. US20020061834A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Roschke, Viktor
; APPLICANT: Li, Yi
; APPLICANT: Ruben, Steven, M.
; TITLE OF INVENTION: Human G-protein Chemokine Receptor (CCR5) HDGMR10
; FILE REFERENCE: 1488.115000C
; CURRENT APPLICATION NUMBER: US/09/779.880A
; CURRENT FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/181,258
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: US 60/187,999
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: US 60/234,336
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 22
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-779-880A-22

Query Match 42.9%; Score 379; DB 9; Length 352;
Best Local Similarity 35.0%; Pred. No. 1.5e-29;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;
QY 20 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 49
DB 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIQFVGNMLVILLINCKR 60
QY 50 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 83
DB 61 LKSMTDIYLLNLAISDLFFLLTVPFWAHAAQWDFGNTMC--QLLTGLYFIQFSGIFF 118
QY 84 -----LCTRSQEGHLHYTC 97
DB 119 ILLTIDRYLAVHVAFAKARTVTFGVTSVITWVAVFASLPGLIFTRSQEGLHYTC 178
QY 98 SSHFPYSQYQFQKQFQTLKI-----HORVHGG----- 124
DB 179 SSHFPYSQYQFQKQFQTLKIIVILGLVPLVWVICYSGLKTLRCRNEKKRHRAVRLIF 238
QY 125 -----GSKYKGLC-----QFFGLNCCSSNRDLGQHV 154
DB 239 TIMIVYFLWAPYNIIVLLNTFQFFGLNCCSSNRDLQAMQV 281

RESULT 7

US-09-813-653-15
; Sequence 15, Application US/09813653
; Patent No. US20020064770A1
; GENERAL INFORMATION:
; APPLICANT: Nestor, John
; APPLICANT: Wilson, Carol
; APPLICANT: See, Raymond
; APPLICANT: Tan Hehix, Christina
; TITLE OF INVENTION: Binding Compounds and Methods For Identifying Binding Compounds
; FILE REFERENCE: CNS-005
; CURRENT APPLICATION NUMBER: US/09/813.653
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US 60/190,946
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/190,996
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/191,299
; PRIOR FILING DATE: 2000-03-21
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-813-653-15

Query Match 42.9%; Score 379; DB 9; Length 352;
Best Local Similarity 35.0%; Pred. No. 1.5e-29;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;
QY 20 MDYQVSSPIYDINYYTSEPCQKINVKQIAA----- 49
DB 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIQFVGNMLVILLINCKR 60
QY 50 -----YKCGLC-----AAAQWDFGNTMCQHQRVHGHHSYKCG--- 83
DB 61 LKSMTDIYLLNLAISDLFFLLTVPFWAHAAQWDFGNTMC--QLLTGLYFIQFSGIFF 118
QY 84 -----LCTRSQEGHLHYTC 97
DB 119 ILLTIDRYLAVHVAFAKARTVTFGVTSVITWVAVFASLPGLIFTRSQEGLHYTC 178

```
98 SSHFPYSQYQWKNFQTLKI-----HORVHGG-----124
|||||
179 SSHFPYSQYQWKNFQTLKIIVILGLVPLLVWVICYSGLKTLRCRNEKCRHRAVELIF 238
|||||
125 -----GGSYKCGLC-----QEFFGLNCCSSNRLDGHORV 154
|||||
239 TIMIVYFLWAPYINIVLLNTFQEFFGLNCCSSNRLDQAMQV 281
|||||

RESULT 8
JS-09-813-653-17
Sequence 17, Application US/09813653
Patent No. US20020064770A1
GENERAL INFORMATION:
APPLICANT: Nestor, John
APPLICANT: Wilson, Carol
APPLICANT: See, Raymond
APPLICANT: Tan Hehir, Christina
TITLE OF INVENTION: Binding Compounds and Methods For Identifying Binding Compounds
FILE REFERENCE: CNS-005
CURRENT APPLICATION NUMBER: US/09/813,653
CURRENT FILING DATE: 2001-03-20
PRIOR APPLICATION NUMBER: US 60/190,946
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: US 60/190,996
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: US 60/191,299
PRIOR FILING DATE: 2000-03-21
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn version 3.0
SEQ ID NO 17
LENGTH: 352
TYPE: PRT
ORGANISM: Homo sapiens
JS-09-813-653-17

Query Match 42.9%; Score 379; DB 9; Length 352;
Best Local Similarity 35.0%; Pred. No. 1.5e-29;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

20 MDYQSSPIYDINYTSEPCQKINVKQIAA-----49
|||||
1 MDYQSSPIYDINYTSEPCQKINVKQIAARLLPPLYSLVFIFGVGNMVLILINCKR 60
|||||
50 -----YKCGLC-----AAQWDFGNTMCOHORVGHGHHHSYKCG---83
|||||
61 LKSMTDIYLLNLAIISDLFLLTVFPWAHYAAQWDFGNTMC--QLLTGLYFIFGFSGIF 118
|||||
84 -----LCTRSQKEGLHYTC 97
|||||
119 ILLTIDRYLAVVHAFALKARTVTGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178
|||||
98 SSHFPYSQYQWKNFQTLKI-----HORVHGG-----124
|||||
179 SSHFPYSQYQWKNFQTLKIIVILGLVPLLVWVICYSGLKTLRCRNEKCRHRAVELIF 238
|||||
125 -----GGSYKCGLC-----QEFFGLNCCSSNRLDGHORV 154
|||||
61 LKSMTDIYLLNLAIISDLFLLTVFPWAHYAAQWDFGNTMC--QLLTGLYFIFGFSGIF 118
|||||
84 -----LCTRSQKEGLHYTC 97
|||||
119 ILLTIDRYLAVVHAFALKARTVTGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178
|||||
98 SSHFPYSQYQWKNFQTLKI-----HORVHGG-----124
|||||

RESULT 9
JS-09-796-202-1
Sequence 1, Application US/09796202
Patent No. US20020068813A1
GENERAL INFORMATION:
APPLICANT: Dragic, Tatjana
APPLICANT: Olson, William
TITLE OF INVENTION: SULFATED CCR5 PEPTIDES FOR HIV-1 INFECTION
FILE REFERENCE: 2048/61010/JPW/SHS
CURRENT APPLICATION NUMBER: US/09/796,202
CURRENT FILING DATE: 2001-02-28
```

```
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 352
TYPE: PRT
ORGANISM: human
US-09-796-202-1

Query Match 42.9%; Score 379; DB 9; Length 352;
Best Local Similarity 35.0%; Pred. No. 1.5e-29;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

20 MDYQSSPIYDINYTSEPCQKINVKQIAA-----49
|||||
1 MDYQSSPIYDINYTSEPCQKINVKQIAARLLPPLYSLVFIFGVGNMVLILINCKR 60
|||||
50 -----YKCGLC-----AAQWDFGNTMCOHORVGHGHHHSYKCG---83
|||||
61 LKSMTDIYLLNLAIISDLFLLTVFPWAHYAAQWDFGNTMC--QLLTGLYFIFGFSGIF 118
|||||
84 -----LCTRSQKEGLHYTC 97
|||||
119 ILLTIDRYLAVVHAFALKARTVTGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178
|||||
98 SSHFPYSQYQWKNFQTLKI-----HORVHGG-----124
|||||
179 SSHFPYSQYQWKNFQTLKIIVILGLVPLLVWVICYSGLKTLRCRNEKCRHRAVELIF 238
|||||
125 -----GGSYKCGLC-----QEFFGLNCCSSNRLDGHORV 154
|||||
239 TIMIVYFLWAPYINIVLLNTFQEFFGLNCCSSNRLDQAMQV 281
|||||

RESULT 10
US-09-195-662A-2
Sequence 2, Application US/09195662A
Patent No. US20020076745A1
GENERAL INFORMATION:
APPLICANT: Li, Yi
APPLICANT: Ruben, Steven, M.
TITLE OF INVENTION: Human G-Protein Chemokine Receptor HDGNR10 (CCR5 Receptor)
FILE REFERENCE: 1488.1150002
CURRENT APPLICATION NUMBER: US/09/195,662A
CURRENT FILING DATE: 1998-11-18
PRIOR APPLICATION NUMBER: 08/466,343
PRIOR FILING DATE: 1995-06-06
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 352
TYPE: PRT
ORGANISM: Artificial Sequence: Genomic
FEATURE:
OTHER INFORMATION: Deduced Amino Acid Sequence
US-09-195-662A-2

Query Match 42.9%; Score 379; DB 9; Length 352;
Best Local Similarity 35.0%; Pred. No. 1.5e-29;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

20 MDYQSSPIYDINYTSEPCQKINVKQIAA-----49
|||||
1 MDYQSSPIYDINYTSEPCQKINVKQIAARLLPPLYSLVFIFGVGNMVLILINCKR 60
|||||
50 -----YKCGLC-----AAQWDFGNTMCOHORVGHGHHHSYKCG---83
|||||
61 LKSMTDIYLLNLAIISDLFLLTVFPWAHYAAQWDFGNTMC--QLLTGLYFIFGFSGIF 118
|||||
84 -----LCTRSQKEGLHYTC 97
|||||
119 ILLTIDRYLAVVHAFALKARTVTGVVTSVITWVAVFASLPGIIFTRSQKEGLHYTC 178
|||||
98 SSHFPYSQYQWKNFQTLKI-----HORVHGG-----124
|||||
```

```
Db 179 SSSHPYSQYQFWKQFOTLKIVILGLVPLLVVICYSGLKTLRCRNEKRRHRAVLIF 238
QY 125 -----GGSYKCGLC-----QEFFGLNCCSSNRLDGHQV 154
Db 239 TIMIVYFLWAPYNIIVLLNTFQEFFGLNCCSSNRLDQAMQV 281

RESULT 11
US-09-339-912A-2
; Sequence 2, Application US/09339912A
; Patent No. US20020099176A1
; GENERAL INFORMATION:
; APPLICANT: Li, Yi
; APPLICANT: Ruben, Steven, M.
; TITLE OF INVENTION: Antibodies to Human G-Protein Chemokine Receptor HDGMR10
; FILE OF INVENTION: (CCRS Receptor)
; FILE REFERENCE: 1488.1150003
; CURRENT APPLICATION NUMBER: US/09/339,912A
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 09/195,662
; PRIOR FILING DATE: 1998-11-18
; PRIOR APPLICATION NUMBER: 08/466,343
; PRIOR FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 2
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Artificial Sequence: Genomic
; FEATURE:
; OTHER INFORMATION: Deduced Amino Acid Sequence
US-09-339-912A-2

Query Match 42.9%; Score 379; DB 9; Length 352;
Best Local Similarity 35.0%; Pred. No. 1.5e-29;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

QY 20 MDYQSSPIYDINYTSPCKINVKQIAA-----49
Db 1 MDYQSSPIYDINYTSPCKINVKQIAARLLPPLYSLVIFGVGNMLVILINCKR 60
QY 50 -----YKCGLC-----AAAQWDFGNTMCQHORVHGHHSYKCG---83
Db 61 LKSMTDIYLLNLALSDFLLTVPFWAHYAAQWDFGNTMC--QLLTGLYFIFGFSGIFP 118
QY 84 -----LCTRSQKEGLHYTC 97
Db 119 ILLTIDRYLAVVHAFKARTVTVGVVTSVITWVAVFASLPGLIIFTRSQKEGLHYTC 178
QY 98 SSSHPYSQYQFWKQFOTLKI-----LCTRSQKEGLHYTC 97
Db 179 SSSHPYSQYQFWKQFOTLKIVILGLVPLLVVICYSGLKTLRCRNEKRRHRAVLIF 238
QY 125 -----GGSYKCGLC-----QEFFGLNCCSSNRLDGHQV 154
Db 239 TIMIVYFLWAPYNIIVLLNTFQEFFGLNCCSSNRLDQAMQV 281

RESULT 12
US-09-938-719-5
; Sequence 5, Application US/09938719
; Patent No. US20020106742A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; APPLICANT: PARMENTIER, MARC
; APPLICANT: VASSART, GILBERT
; APPLICANT: LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
```

```
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: U.S.A.
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/938,719
FILING DATE: 24-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/626,939
FILING DATE: 27-JULY-2000
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER: <Unknown>
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-938-719-5

Query Match 42.9%; Score 379; DB 10; Length 352;
Best Local Similarity 35.0%; Pred. No. 1.5e-29;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

QY 20 MDYQSSPIYDINYTSPCKINVKQIAA-----49
Db 1 MDYQSSPIYDINYTSPCKINVKQIAARLLPPLYSLVIFGVGNMLVILINCKR 60
QY 50 -----YKCGLC-----AAAQWDFGNTMCQHORVHGHHSYKCG---83
Db 61 LKSMTDIYLLNLALSDFLLTVPFWAHYAAQWDFGNTMC--QLLTGLYFIFGFSGIFP 118
QY 84 -----LCTRSQKEGLHYTC 97
Db 119 ILLTIDRYLAVVHAFKARTVTVGVVTSVITWVAVFASLPGLIIFTRSQKEGLHYTC 178
QY 98 SSSHPYSQYQFWKQFOTLKI-----LCTRSQKEGLHYTC 97
Db 179 SSSHPYSQYQFWKQFOTLKIVILGLVPLLVVICYSGLKTLRCRNEKRRHRAVLIF 238
QY 125 -----GGSYKCGLC-----QEFFGLNCCSSNRLDGHQV 154
Db 239 TIMIVYFLWAPYNIIVLLNTFQEFFGLNCCSSNRLDQAMQV 281

RESULT 13
US-09-939-226-5
; Sequence 5, Application US/09939226
; Patent No. US20020110805A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; APPLICANT: PARMENTIER, MARC
; APPLICANT: VASSART, GILBERT
; APPLICANT: LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
```

```
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/938,703
  FILING DATE: 24-Aug-2001
  CLASSIFICATION: <Unknown>
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US/09/939,226
    FILING DATE: 24-Aug-2001
    CLASSIFICATION: <Unknown>
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 09/626,939
    FILING DATE: 2000-07-27
    ATTORNEY/AGENT INFORMATION:
      NAME: Altman, Daniel E
      REGISTRATION NUMBER: 34,115
      REFERENCE/DOCKET NUMBER: <Unknown>
    INFORMATION FOR SEQ ID NO: 5:
      SEQUENCE CHARACTERISTICS:
        LENGTH: 352 amino acids
        TYPE: amino acid
        TOPOLOGY: linear
      MOLECULE TYPE: protein
      SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-938-703-5

Query Match      42.9%; Score 379; DB 10; Length 352;
Best Local Similarity 35.0%; Pred. No. 1.5e-29;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

QY 20 MDYQVSSPIYDINYTSPCKINVKQIAA----- 49
Db 1 MDYQVSSPIYDINYTSPCKINVKQIAARLLPLYSLVLPFGVGNMLVILLINCKR 60
QY 50 -----YKCGLC-----AAAQWDFGNTMCOHQRVHGHHSYKCG--- 83
Db 61 LKSWTDIYLLNLAISDLFFLLTPFMAHYAAQWDFGNTMC--QLLTGLYFIFGFSGIF 118
QY 84 -----LCTRSQKEGLHYTC 97
Db 119 IILLTDIYLVAVHAFALKARTVTFGVVTSVITWVAVFASLPGLIFTRSQKEGLHYTC 178
QY 98 SSHFPYQYQFQWKNFQTLKI-----HORVHGG----- 124
Db 179 SSHFPYQYQFQWKNFQTLKIVILGLVLPVLMVVICYSIGILKTLRCRNEKKRHRAVRLIF 238
QY 125 -----GGSYKCGLC-----QEFFGLNCCSSNRLDGHQRV 154
Db 239 TIMIVYFLFWAPYNIYVLLNTFQEFFGLNCCSSNRLDQAMQV 281

RESULT 15
US-09-502-783A-2
; Sequence 2, Application US/09502783A
; Patent No. US20020132269A1
; GENERAL INFORMATION:
; APPLICANT: Li, Yi
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein Chemokine Receptor (CCRS)
; FILE REFERENCE: 1488.1150006
; CURRENT APPLICATION NUMBER: US/09/502,783A
; PRIOR FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: 08/466,343
; PRIOR FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-502-783A-2

Query Match      42.9%; Score 379; DB 10; Length 352;
Best Local Similarity 35.0%; Pred. No. 1.5e-29;
Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

QY 20 MDYQVSSPIYDINYTSPCKINVKQIAA----- 49
Db 1 MDYQVSSPIYDINYTSPCKINVKQIAARLLPLYSLVLPFGVGNMLVILLINCKR 60
QY 50 -----YKCGLC-----AAAQWDFGNTMCOHQRVHGHHSYKCG--- 83
Db 61 LKSWTDIYLLNLAISDLFFLLTPFMAHYAAQWDFGNTMC--QLLTGLYFIFGFSGIF 118
QY 84 -----LCTRSQKEGLHYTC 97
Db 119 IILLTDIYLVAVHAFALKARTVTFGVVTSVITWVAVFASLPGLIFTRSQKEGLHYTC 178
QY 98 SSHFPYQYQFQWKNFQTLKI-----HORVHGG----- 124
Db 179 SSHFPYQYQFQWKNFQTLKIVILGLVLPVLMVVICYSIGILKTLRCRNEKKRHRAVRLIF 238
QY 125 -----GGSYKCGLC-----QEFFGLNCCSSNRLDGHQRV 154
Db 239 TIMIVYFLFWAPYNIYVLLNTFQEFFGLNCCSSNRLDQAMQV 281

RESULT 14
US-09-938-703-5
; Sequence 5, Application US/09938703
; Patent No. US20020110870A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; APPLICANT: FARMETIER, MARC
; APPLICANT: VASSART, GILBERT
; APPLICANT: LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobb, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 17:49:04 ; Search time 21.8203 Seconds
(without alignments)
304.432 Million cell updates/sec

Title: US-10-057-890A-31
Perfect score: 884
Sequence: 1 MKVSVAAALSLMLVTALGSM.....GLNKCSSNRLDGHQVRHAA 157

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCRTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	379	42.9	352	3	US-08-087-232A-13
2	379	42.9	352	3	US-08-861-105-14
3	379	42.9	352	3	US-08-575-967A-2
4	379	42.9	352	4	US-08-833-752-5
5	379	42.9	352	4	US-09-502-783A-2
6	379	42.9	352	4	US-09-796-202-1
7	374	42.3	352	3	US-09-045-583-52
8	374	42.3	352	4	US-09-534-185-52
9	373	42.2	352	3	US-08-466-343D-2
10	359	40.6	352	4	US-09-517-605-5
11	272	30.8	354	4	US-08-724-984A-2
12	258	29.2	184	4	US-08-833-752-4
13	258	29.2	215	3	US-09-087-232A-17
14	258	29.2	215	4	US-08-833-752-6
15	182.5	20.6	100	3	US-09-087-232A-15
16	159	18.0	87	3	US-09-087-232A-18
17	148.5	16.8	488	2	US-08-933-750C-17
18	148.5	16.8	488	3	US-09-234-613-17
19	128	14.5	54	4	US-08-833-752-11
20	123	13.9	22	3	US-08-861-105-4
21	121	13.7	711	2	US-08-820-170A-10
22	121	13.7	711	3	US-09-055-699-10
23	121	13.7	711	3	US-09-273-565-10
24	121	13.7	711	4	US-09-565-538-10
25	121	13.7	711	4	US-09-661-468-10
26	121	13.7	711	4	US-09-976-165-10
27	118.5	13.4	315	1	US-08-253-155A-34

28	118.5	13.4	532	4	US-09-389-956-80	Sequence 80, Appl
29	117	13.2	457	4	US-09-389-956-68	Sequence 68, Appl
30	117	13.2	647	4	US-09-389-956-6	Sequence 6, Appl
31	114.5	13.0	543	4	US-09-362-123A-4	Sequence 4, Appl
32	114.5	13.0	675	1	US-08-317-522A-9	Sequence 9, Appl
33	114.5	13.0	675	1	US-08-439-818A-9	Sequence 9, Appl
34	114.5	13.0	675	2	US-08-751-965-9	Sequence 9, Appl
35	114.5	13.0	675	2	US-08-738-975-9	Sequence 9, Appl
36	114.5	13.0	675	2	US-08-728-626-9	Sequence 9, Appl
37	114.5	13.0	675	3	US-08-808-599A-9	Sequence 9, Appl
38	114	12.9	21	3	US-08-907-468-11	Sequence 11, Appl
39	113.5	12.8	462	3	US-08-486-099-117	Sequence 117, Appl
40	113.5	12.8	462	3	US-08-360-107A-127	Sequence 127, Appl
41	113.5	12.8	462	3	US-08-484-223B-117	Sequence 117, Appl
42	113.5	12.8	462	3	US-08-919-597-117	Sequence 117, Appl
43	113.5	12.8	462	3	US-08-475-668A-117	Sequence 117, Appl
44	113.5	12.8	462	3	US-08-485-551A-117	Sequence 117, Appl
45	113.5	12.8	462	3	US-08-471-913A-117	Sequence 117, Appl

ALIGNMENTS

RESULT 1
US-09-087-232A-13
; Sequence 13, Application US/09087232A
; Patent No. 6153431
; GENERAL INFORMATION:
; APPLICANT: Quillent et al.
; TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS CO-RECEPTOR
; TITLE OF INVENTION: VARIANTS ASSOCIATED WITH RESISTANCE TO VIRUS INFECTION.
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Baker & Botts, L.L.P. attn. Lisa Kole
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION NUMBER: US/09/087.232A
FILING DATE: 28 MAY 1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/048,057
FILING DATE: 30 MAY 1997
ATTORNEY/AGENT INFORMATION:
NAMES: KOLE, LISA B.
REGISTRATION NUMBER: 35,225
REFERENCE/DOCKET NUMBER: AP 31115
TELEPHONE: (212) 408-2628
TELEFAX: (212) 765-2519
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

Query Match 42.9%; Score 379; DB 3; Length 352;

Best Local Similarity 35.0%; Pred. No. 7e-35;

Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

Qy 20 MDYQVSSPIYDINTYTTSEPCOKINVKQIAA-----49

Db 1 MDYQVSSPIYDINTYTTSEPCOKINVKQIAARLLPPLYSLVFIFGVGNMLVILINCKR 60

bb 1 MDYQVSSPIYDINYTSEPCOKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKR 60
dy 50 -----YKGLC-----AAAQWDFGNTM 66
bb 61 LKSMTDIYLLNLAIISDLFFLLTVPFWAHYAAAQWDFGNTM 100

Search completed: October 21, 2003, 17:55:25
Job time : 23.8203 secs

CC (e.g. arthritis, Graves disease, multiple sclerosis (MS), autoimmune
 CC myocarditis, diabetes and lupus) by administering a peptide construct
 CC comprising a TCBL peptide bonded to an antigenic peptide associated with
 CC the disorder. Unlike prior art peptide conjugates, a modified version of
 CC peptide G has long range stabilisation and also enhances the immune
 CC response. AAU82019-AAU82114 represent T-cell specific binding ligand
 CC peptides, peptide constructs or peptides used in their construction.
 XX
 XX Sequence 6 AA;
 SQ

Query Match 100.0%; Score 5; DB 23; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
 Db 1 GGGGS 5
 |||||

RESULT 86
 AAU82019
 ID AAU82019 standard; peptide; 6 AA.
 XX
 AC AAU82019;
 DT 28-MAY-2003 (first entry)
 XX
 DE Peptide linker #2.
 XX

Siglec-12; therapy; organ transplant disorder; diabetes mellitus;
 KW haematologic disorder; Fanconi's aplastic anaemia; septic shock; stroke;
 KW Alzheimer's disease; infection; asthma; cachexia; myocardial infarction;
 KW immunosuppressive; thyromimetic; antibacterial; antiparasitic; cardiant;
 KW liver disorder; cerebroprotective; infection; viricide; neuroprotective;
 KW lymphoma; osteoporosis; dermatological; nootropic; rheumatoid arthritis;
 KW dementia; protozoacide; vasotropic; Hashimoto's thyroiditis; cytostatic;
 KW ischaemia; reperfusion injury; systemic lupus erythematosus; psoriasis;
 KW osteopathic.
 XX
 OS Unidentified.
 XX

Key Location/Qualifiers
 FH Misc-difference 6
 FT /note= "X at this position represents one or more
 FT repeats of GGGGS"
 XX
 XX WO200296452-A1.
 XX
 XX 05-DEC-2002.
 XX
 XX 29-MAY-2002; 2002WO-US16906.
 XX
 XX 29-MAY-2001; 2001US-294199P.
 XX
 XX (IMMV) IMMUNEX CORP.
 XX
 XX Anderson DM, Marken JS;
 XX
 XX WPI; 2003-140419/13.
 XX

Novel substantially purified siglec-12 polypeptide useful for treating
 PT rheumatoid arthritis, systemic lupus erythematosus, Hashimoto's
 PT thyroiditis, asthma, Alzheimer's disease, dementia, cachexia,
 PT myocardial infarction -
 XX
 XX Disclosure; Column 16; 47pp; English.
 XX

The invention relates to Siglec-12 polypeptide and polynucleotide
 CC sequences. Siglec-12 is useful for identifying an agent that modulates
 CC its activity. The invention is useful for treating a Siglec-associated
 CC disorder or disease (e.g. rheumatologic disorder, bone marrow or solid
 CC organ transplant disorder, a graft-versus-host disorder, inflammatory
 CC disorder, autoimmune disorder, neurologic disorder, cell proliferative

CC disorder, infection, cardiovascular disorder, haematologic disorder,
 CC liver disorder or a bone disorder). Antibody that specifically binds
 CC to Siglec-12 is useful for treating a tumour that expresses Siglec-12
 CC polypeptide. The invention is also useful for treating systemic lupus
 CC erythematosus, rheumatoid arthritis, dementia, lymphoma, Hashimoto's
 CC thyroiditis, Alzheimer's disease, bacterial, parasitic, protozoal and
 CC viral infections, asthma, cachexia, myocardial infarction, ischaemia/
 CC reperfusion injury, stroke, diabetes mellitus, psoriasis, Fanconi's
 CC aplastic anaemia, septic shock, osteoporosis, etc. The present sequence
 CC is a peptide linker used in the invention.
 XX
 XX Sequence 6 AA;
 SQ

Query Match 100.0%; Score 5; DB 24; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
 Db 1 GGGGS 5
 |||||

RESULT 87
 ABP71089
 ID ABP71089 standard; peptide; 6 AA.
 XX
 AC ABP71089;
 DT 14-APR-2003 (first entry)
 XX

Amino acid sequence of a peptide linker.
 DE
 XX HAM; homologue of attractin/mahogany; immunosuppressive; cytostatic;
 XX antiinflammatory; cardiant; osteopathic; gene therapy; linker.
 KW
 KW Synthetic.
 XX
 OS
 XX

Key Location/Qualifiers
 FH Misc-difference 6
 FT /note= "Xaa is one or more repeats of GGGGS"
 FT
 FT WO200297120-A1.
 XX
 XX 05-DEC-2002.
 XX
 XX 23-MAY-2002; 2002WO-US16391.
 XX
 XX 25-MAY-2001; 2001US-293608P.
 XX
 XX 24-SEP-2001; 2001US-324626P.
 XX
 XX (IMMV) IMMUNEX CORP.
 XX
 XX Anderson DM;
 XX
 XX WPI; 2003-140486/13.
 XX

New Homologue of Attractin/Mahogany (HAM) polypeptide, useful for
 PT treating HAM-associated disorder consisting of inflammatory, -
 PT autoimmune, cell proliferative or cardiovascular disorders -
 XX
 XX Disclosure; Page 14; 89pp; English.
 XX

The invention relates to Homologue of Attractin/Mahogany (HAM)
 CC polypeptides and encoding polynucleotides. The HAM polypeptides can be
 CC expressed by standard recombinant methodology. The HAM polypeptides are
 CC useful for treating HAM-associated disorder consisting of inflammatory,
 CC autoimmune, graft-versus-host, neurologic, myelination, cell
 CC proliferative, cardiovascular, haematologic, liver, metabolic, weight or
 CC bone disorder. The present sequence represents a peptide linker.
 XX
 XX Sequence 6 AA;
 SQ

Query Match 100.0%; Score 5; DB 24; Length 6;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

AY 1 GGGGS 5
|
|
|
|
|
DB 1 GGGGS 5

RESULT 88
AAR99245
D AAR99245 standard; Peptide; 7 AA.
C AAR99245;
C AAR99245;
DT 28-NOV-1996 (first entry)
C (Gly6)Ser linker.
C Bioactive fusion protein; interleukin-12; IL-12; p35 subunit;
C p40 subunit; antitumor; cytokine; tumor; melanoma;
C fibrosarcoma; renal cell carcinoma; immunotherapy; therapy;
C retrovirus; vector.
C Synthetic.
C WO9624676-A1.
C 15-AUG-1996.
C 07-FEB-1996; 96WO-US01787.
C 08-FEB-1995; 95US-0385335.
C (WHED) WHITEHEAD INST BIOMEDICAL RES.
C Lieschke GJ, Mulligan RC;
C WPI; 1996-384448/38.
C N-PSDB; AAT35197.
C T New DNA encoding fusion protein, esp. contg. IL-12 p35 and p40
T subunits - for treatment of established tumours or prevention of
T tumour establishment
C Claim 2; Page 69; 118pp; English.
C Peptide linker (Gly4Ser)2Ser, (Gly4Ser)3Ser, (Gly4Ser)3 and
C (Gly6)Ser (AAR99242.45) are used to join the subunits of novel
C dimeric or multimeric fusion proteins. They have been utilised in
C the prodn. of bioactive interleukin-12 (IL-12) fusion proteins,
C linking mouse/human IL-12 p35 subunit (see also AAR99246) to
C mouse/human IL-12 p40 subunit (AAR99247). DNA encoding such constructs
C can be incorporated into a retroviral vector (see also AAT35198) to
C allow dimeric IL-12 prodn. in transfected cells. Tumour cells (esp.
C CMS-5, B16 or renal carcinoma cells) secreting IL-12 dimer can be
C used to reduce the size of established tumours and/or increase
C survival time, esp. in cases of melanoma, fibrosarcoma and renal
C cell carcinoma.
C Sequence 7 AA;
Q

Query Match 100.0%; Score 5; DB 17; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

AY 1 GGGGS 5
|
|
|
|
|
b 3 GGGGS 7

RESULT 89
AY23703
D AAY23703 standard; Peptide; 7 AA.

XX AAY23703;
AC 07-SEP-1999 (first entry)
DT Peptide identified using the method of the invention.
DE Polysome library; nascent peptide display;
XX affinity interaction screening; multiple binding specificity;
XX single chain antibody.
KW Synthetic.
OS US5922545-A.
PN 13-JUL-1999.
XX 29-JUL-1997; 97US-0902623.
XX 17-JAN-1996; 96US-0586176.
PR 29-OCT-1993; 93US-0144775.
PR 02-SEP-1994; 94US-0300262.
PR 25-OCT-1994; 94WO-US12206.
PR 29-JUL-1997; 97US-0902623.
XX (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX Dower WJ, Mattheakis LC;
XX WPI; 1999-417975/35.
DR In vitro peptide and antibody display libraries for identifying
PT multiple binding specificities of single chain antibodies
XX Example 1; Column 71-72; 50pp; English.
XX The specification describes a method for generating libraries of
CC polysomes displaying nascent peptides suitable for affinity
CC interaction screening. The method can be used to identify multiple
CC binding specificities of single chain antibodies, and comprises
CC contacting antigens with a polysome library displaying nascent
CC peptides having a single chain antibody segment, separating polysomes
CC bound to the antigen species from polysomes not bound to the antigen
CC species, and synthesising cDNA from the separated bound polysomes to
CC identify single chain antibodies which bind to at least one of the
CC antigen species present. The method for generating libraries of
CC polysomes displaying nascent peptides is useful for affinity interaction
CC screening. The present sequence was identified using the method of the
CC invention.
XX Sequence 7 AA;
SQ Query Match 100.0%; Score 5; DB 20; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGGS 5
|
|
|
|
|
DB 3 GGGGS 7

RESULT 90
AAY02129
ID AAY02129 standard; Protein; 7 AA.
XX AAY02129;
XX 16-JUL-1999 (first entry)
DT Peptide linker used to make multifunctional proteins.
DE Angiostatin; endostatin; interferon; thrombospondin;
XX interferon-inducible protein; platelet factor 4; anti-angiogenic;
KW

CC feline IL-12 single chain or canine IL-12 single chain proteins, a
CC nucleic acid encoding these proteins, mimetopes of these proteins,
CC multimeric forms of these proteins, an antibody against these proteins,
CC or an inhibitor identified by its ability to inhibit the activity of
CC these proteins, can be used to treat or prevent autoimmune diseases,
CC allergic reactions, infectious diseases, tumour development, inflammatory
CC diseases and/or graft rejection in animals. The present amino acid
CC sequence represents a synthetic peptide used in the methods of the
XX invention.
SQ Sequence 7 AA;
Query Match 100.0%; Score 5; DB 23; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGGS 5
DB 1 GGGGS 5

CC the treatment of pathologies associated with botulinum neurotoxin
CC poisoning, for rapid detection/diagnosis of botulism and in the detection
CC and/or quantification of BONT/A in a biological sample obtained from an
CC organism which is indicative of a Clostridium botulinum infection of
CC the organism. The present sequence is a linker peptide which may be
CC incorporated into a single chain antibody (scFv) of the invention.
XX
SQ Sequence 7 AA;
Query Match 100.0%; Score 5; DB 24; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGGS 5
DB 1 GGGGS 5

X X X X
X X X X

Example 1: Page 19, JEPF, Row 14b.

[illegible]

TELEFAX: (213) 977-1003
; INFORMATION FOR SEQ ID NO: 76:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US94-02631-76

Query Match 100.0%; Score 5; DB 5; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
DB 2 GGGGS 6

RESULT 89
US-08-902-623-49
; Sequence 49, Application US/08902623
; Patent No. 5922545
; GENERAL INFORMATION:
; APPLICANT: MATTHEAKIS, LARRY C.
; APPLICANT: DOWER, WILLIAM J.
; TITLE OF INVENTION: IN VITRO PEPTIDE AND ANTIBODY DISPLAY
; TITLE OF INVENTION: LIBRARIES
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND AND TOWNSEND AND CREW LLP
; STREET: TWO EMBARCADERO CENTER, 8TH FLOOR
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: U.S.A.
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/902.623
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/586.176
; FILING DATE: 17-JAN-1996
; APPLICATION NUMBER: US 08/300.262
; FILING DATE: 02-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/144.775
; FILING DATE: US 29-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/12206
; FILING DATE: US 25-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: DUNN, TRACY J.
; REGISTRATION NUMBER: 34,587
; REFERENCE/DOCKET NUMBER: 16528X-003230US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
US-08-902-623-49

Query Match 100.0%; Score 5; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
DB 3 GGGGS 7

RESULT 90
US-09-374-454-16
; Sequence 16, Application US/09374454
; Patent No. 6395548
; GENERAL INFORMATION:
; APPLICANT: Lee, Mu-En
; APPLICANT: Maemura, Koji
; APPLICANT: Hsieh, Chung-Wing
; TITLE OF INVENTION: METHODS OF MODULATING OF ANGIOGENESIS
; FILE REFERENCE: 05433/037001
; CURRENT APPLICATION NUMBER: US/09/374.454
; CURRENT FILING DATE: 1999-08-13
; EARLIER APPLICATION NUMBER: US 60/096.515
; EARLIER FILING DATE: 1998-08-14
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated peptide
US-09-374-454-16

Query Match 100.0%; Score 5; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
DB 1 GGGGS 5

RESULT 91
US-08-897-956A-38
; Sequence 38, Application US/08897956A
; Patent No. 6423512
; GENERAL INFORMATION:
; APPLICANT: Mary Ellen Digan
; APPLICANT: Philip Lake
; APPLICANT: Hermann Gram
; TITLE OF INVENTION: Fusion Polypeptides
; FILE REFERENCE: 600-7244/CPA
; CURRENT APPLICATION NUMBER: US/08/897.956A
; CURRENT FILING DATE: 1997-07-21
; PRIOR APPLICATION NUMBER: 60/022.689
; PRIOR FILING DATE: 1996-07-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Linker
US-08-897-956A-38

Query Match 100.0%; Score 5; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
DB 3 GGGGS 7

RESULT 92

```

JS-08-750-128-6
Sequence 6, Application US/08750128
Patent No. 5916773
GENERAL INFORMATION:
APPLICANT: MELE Antonio,
APPLICANT: DE SANTIS Rita,
APPLICANT: CARLONI Cristina,
APPLICANT: COSCARELLA Annamaria
TITLE OF INVENTION: Hybrid molecule of formula GM-CSF-L-BPO or
TITLE OF INVENTION: EPO-L-GM-CSF for hematopoietic stimulation
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: LOWE PRICE, LeBLANC & BECKER
STREET: 99 Canal Center Plaza, Suite 300
CITY: Alexandria
STATE: VIRGINIA
COUNTRY: USA
ZIP: 22314
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk, 3.50 inch.
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/750,128
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IT FI 94 A 000106
FILING DATE: 27-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Robert L. Price
REGISTRATION NUMBER: 22,685
REFERENCE/DOCKET NUMBER: 2879-007
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-684-1111
TELEFAX: 703-684-1124
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
JS-08-750-128-6
Query Match 100.0%; Score 5; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2y 1 GGGGS 5
Db 4 GGGGS 8

RESULT 93
JS-08-487-890A-144
Sequence 144, Application US/08487890A
Patent No. 5708149
GENERAL INFORMATION:
APPLICANT: Loosmore, Sheena
APPLICANT: Harkness, Robin
APPLICANT: Schryvers, Anthony
APPLICANT: Chong, Pele
APPLICANT: Gray-Owen, Scott
APPLICANT: Yang, Yan-Ping
APPLICANT: Murdin, Andrew
APPLICANT: Klein, Michel
TITLE OF INVENTION: Transferrin Receptor Genes
NUMBER OF SEQUENCES: 147
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,890A

```

```

; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,890A
; FILING DATE: 07-JUN-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/175,116
; FILING DATE: 29-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,968
; FILING DATE: 08-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-466 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 144:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-487-890A-144
Query Match 100.0%; Score 5; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5
Db 3 GGGGS 7

RESULT 94
US-08-487-890A-145
Sequence 145, Application US/08487890A
Patent No. 5708149
GENERAL INFORMATION:
APPLICANT: Loosmore, Sheena
APPLICANT: Harkness, Robin
APPLICANT: Schryvers, Anthony
APPLICANT: Chong, Pele
APPLICANT: Gray-Owen, Scott
APPLICANT: Yang, Yan-Ping
APPLICANT: Murdin, Andrew
APPLICANT: Klein, Michel
TITLE OF INVENTION: Transferrin Receptor Genes
NUMBER OF SEQUENCES: 147
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,890A

```

```

XX DE GM-CSF/EPO linker fragment, L2.
XX KW Linker sequence; granulocyte macrophage-colony stimulating factor;
XX KW GM-CSF; erythropoietin; EPO; hybrid protein; stimulation;
XX KW haematopoiesis; erythroid differentiation.
XX OS Synthetic.
XX PN WO9533057-A1.
XX PD 07-DEC-1995.
XX PF 26-MAY-1995; 95WO-EP02011.
XX PR 27-MAY-1994; 94IT-FI00106.
XX PA (MENA ) MENARINI RICERCHE SUD SPA.
XX PI Carloni C, Coscarella A, De Santis R, Mele A;
XX XX WPI; 1996-030568/03.
XX DR N-PSDB; AAT06970.
XX PT GM-CSF-EPO hybrid proteins contg. linker sequence - to stimulate
XX PT haematopoiesis, with higher action specificity compared to unlinked
XX PT granulocyte macrophage colony stimulating factor (GM-CSF) and
XX PT erythropoietin (EPO)
XX PS Claim 7; Fig 3; 26pp; English.
XX XX The sequences given in AAR86794-96 represent the linker sequences of the
XX CC invention. These linkers are used to join granulocyte macrophage-
XX CC colony stimulating factor (GM-CSF) and erythropoietin (EPO) in the
XX CC formation of a hybrid protein. The hybrid protein is useful for the
XX CC stimulation of haematopoiesis. The fused molecules exhibit a higher
XX CC specificity of action on erythroid differentiation, compared to that
XX CC of an equimolar mixture of unlinked GM-CSF and EPO molecules.
XX CC (Updated on 25-MAR-2003 to correct PR field.)
XX SQ Sequence 8 AA;
XX Query Match 100.0%; Score 5; DB 17; Length 8;
XX Best Local Similarity 100.0%; Pred. No. 9.3e+05;
XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX Qy 1 GGGGS 5
XX Db ||||| 8
XX 4 GGGGS 8
RESULT 100
AAY43498
ID AAY43498 standard; Peptide; 8 AA.
XX AC AAY43498;
XX DT 26-JAN-2000 (first entry)
XX DE Linker for dual avb3 receptor/metastasis-associated receptor ligands.
XX KW Interferon-alpha-2b; IFN-alpha; avb3 antagonist; avb3 receptor ligand;
XX KW metastasis-associated receptor ligand; angiogenesis; cell proliferation;
XX KW anti-angiogenic protein; avb3-integrin; cancer; arthritis;
XX KW macular degeneration; diabetic retinopathy; hemangioma; psoriasis;
XX KW osteoporosis; thrombosis; angina; atherosclerosis; antiviral;
XX KW antibacterial; antifungal.
XX OS Homo sapiens.
XX PN WO9951638-A1.
XX XX 14-OCT-1999.
XX PF 07-APR-1999; 99WO-US04295.
XX PR 08-APR-1998; 98US-0081074.
XX PA (SEAR ) SEARLE & CO G D.
XX PI Tjoeng FS, Fok KP;
XX DR WPI; 1999-620196/53.
XX PT New conjugates of integrin antagonist and ligand for
XX PT metastasis-associated receptor, for treating angiogenesis-related
XX PT diseases, e.g. cancer
XX PS Claim 18; Page 86; 108pp; English.
XX XX The present sequence represents a linker used to join the avb3
XX CC antagonist and the metastasis-associated receptor ligand, in the
XX CC pharmaceutical compounds of the invention. These compounds are dual
XX CC avb3 receptor/metastasis-associated receptor ligands, and inhibit
XX CC angiogenesis and thus proliferation of (cancer) cells. One component
XX CC binds to the avb3 receptor and the other to a metastasis-associated
XX CC receptor. The avb3 antagonists may also be conjugated to anti-angiogenic
XX CC proteins, such as IFN-alpha and its derivatives. The compounds are used
XX CC to treat angiogenesis-related disorders (mediated by the avb3-integrin),
XX CC specifically cancer (of lung, breast, ovary, prostate, stomach, colon,
XX CC kidney or bladder, also melanoma, hepatoma, sarcoma and lymphoma),
XX CC arthritis and macular degeneration, and also diabetic retinopathy,
XX CC hemangioma, psoriasis, osteoporosis, thrombosis, angina, atherosclerosis
XX CC etc. The compounds may also be useful as antiviral, antibacterial and
XX CC antifungal agents.
XX SQ Sequence 8 AA;
XX Query Match 100.0%; Score 5; DB 20; Length 8;
XX Best Local Similarity 100.0%; Pred. No. 9.3e+05;
XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX Qy 1 GGGGS 5
XX Db ||||| 8
XX 4 GGGGS 8
Search completed: October 21, 2003, 18:23:57
Job time : 45.7778 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 17:47:49 ; Search time 26.078 Seconds
(without alignments)
578.975 Million cell updates/sec

Title: US-10-057-890A-31

Perfect score: 884

Sequence: 1 MKVSVAAALSCMLVLTALGSM.....GLNNCSSNRLDGHQRVHAA 157

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_76:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	379	42.9	352	2 A43113	chemokine (C-C) re
2	154	17.4	201	2 I57505	zinc finger protei
3	144	16.3	693	2 I37570	zinc finger protei
4	143.5	16.2	319	2 T46469	hypothetical prote
5	141	16.0	555	2 I53869	zinc finger protei
6	138.5	15.7	223	2 F31201	GLI-related finger
7	136.5	15.4	348	2 I38599	zinc finger protei
8	136.5	15.4	386	2 T12527	hypothetical prote
9	135.5	15.3	572	2 I39311	Kruppel-type zinc
10	134	15.2	710	2 I48668	zinc finger protei
11	133.5	15.1	475	2 S03679	finger protein (cl
12	133.5	15.1	636	2 I48689	gene NK10 protein
13	133.5	15.1	701	2 T14757	hypothetical prote
14	131.5	14.9	686	2 A34612	zinc finger protei
15	131	14.8	261	2 S70006	finger protein xfo
16	129.5	14.6	614	2 JH0500	zinc finger protei
17	128.5	14.5	580	2 A37107	spermatogenesis pr
18	127.5	14.4	469	2 I38600	zinc finger protei
19	127.5	14.4	1173	2 I50620	procr2 - chicken
20	127	14.4	728	2 A48830	probable transcrip
21	126.5	14.3	732	2 S47073	finger protein HZF
22	124	14.0	367	2 S06582	finger protein (cl
23	124	14.0	589	2 I38598	zinc finger protei
24	123.5	14.0	209	2 S47068	finger protein HZF
25	123.5	14.0	347	2 S00549	developmental cont
26	123.5	14.0	378	2 S33994	finger protein ZNF
27	122.5	13.9	488	2 S47072	finger protein HZF
28	122.5	13.9	594	2 T12488	hypothetical prote
29	122.5	13.9	803	2 S26823	zinc finger protei

30 122 13.8 1350 2 S00647
31 121.5 13.7 337 2 S60520
32 120.5 13.6 169 2 A39240
33 120.5 13.6 229 2 A48927
34 120.5 13.6 292 2 S43826
35 120.5 13.6 399 2 S47071
36 120.5 13.6 546 2 I49636
37 120.5 13.6 1191 2 S35305
38 120 13.6 435 2 S00833
39 120 13.6 654 2 A57885
40 119.5 13.5 194 2 I53859
41 119.5 13.5 393 2 JN0533
42 119.5 13.5 576 2 A48157
43 119.5 13.5 671 2 JF0288
44 118.5 13.4 107 2 I39315
45 118.5 13.4 120 2 G02493

finger protein - A
finger protein ZNF
finger protein mfg
kruppel-like zinc
finger protein OZF
DNA-binding protei
zinc finger protei
finger protein (cl
finger protein ZNF
zinc finger protei
finger protein PML
renal transcriptio
kruppel-type zinc
zinc-finger protei
KR-ZNF1 - human (f

ALIGNMENTS

RESULT 1

A43113
chemokine (C-C) receptor 5 - human
N; Alternate names: C-C CKR-5; CCR5
C; Species: Homo sapiens (man)
C; Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 20-Jun-2000
C; Accession: A43113; S71808; A58834; A58832; G02653; A58833
R; Samson, M.; Labbe, O.; Mollereau, C.; Vassart, G.; Parmentier, M.
Biochemistry 35, 3362-3367, 1996
A; Title: Molecular cloning and functional expression of a new human CC-chemokine recept
A; Reference number: A43113; MUID:96241590; PMID:8639485
A; Accession: A43113
A; Molecule type: mRNA
A; Residues: 1-352 <SAML>
A; Cross-references: GB:X91492; NID:G1362810; PIDN:CAA62796.1; PID:G1262811
R; Samson, M.; Libert, F.; Doranz, B.J.; Rucker, J.; Liesnard, C.; Farber, C.M.; Saragol
M.; Inai, T.; Rana, S.; Yi, Y.; Smyth, R.J.; Collman, R.G.; Doms, R.W.; Vassart, G.; f
Nature 382, 722-725, 1996
A; Title: Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles
A; Reference number: S71808; MUID:96345670; PMID:8751444
A; Accession: S71808
A; Status: nucleic acid sequence not shown; not compared with conceptual translation
A; Molecule type: DNA
A; Residues: 182-206; 207-230 <SAM2>
A; Accession: A58834
A; Status: nucleic acid sequence not shown; not compared with conceptual translation
A; Molecule type: DNA
A; Residues: 1-184, 'IKSHLGNPAAACHGHLLGNPKNSASVSK' <SAM3>
A; Cross-references: GB:X99393; NID:G1524062; PIDN:CAA67767.1; PID:G1524063
A; Note: this frameshift mutation results in a non-functional receptor but confers a deg
nd may have had a selective advantage by conferring resistance to Yersinia plague infec
R; Combadiere, C.; Ahuja, S.K.; Tiffany, H.L.; Murphy, P.M.
J. Leukoc. Biol. 60, 147-152, 1996
A; Title: Cloning and functional expression of CC CKR5, a human monocyte CC chemokine re
A; Reference number: A58833; MUID:96295970; PMID:8699119
A; Accession: A58832
A; Molecule type: mRNA
A; Residues: 1-352 <COM1>
A; Cross-references: GB:U57840; NID:G1502408; PIDN:AAB17071.1; PID:G1502409
A; Experimental source: clone 8, endotoxin-stimulated peripheral blood monocytes
R; Combadiere, C.
submitted to the EMBL Data Library, May 1996
A; Reference number: H01541
A; Accession: G02653
A; Status: translated from GB/EMBL/DBDJ
A; Molecule type: mRNA
A; Residues: 1-89, 'L', '91-352 <COM2>
A; Cross-references: EMBL:U57840
R; Raport, C.J.; Gosling, J.; Schweickart, V.L.; Gray, P.W.; Charo, I.F.
J. Biol. Chem. 271, 17161-17166, 1996
A; Title: Molecular cloning and functional characterization of a novel human CC chemokin
A; Reference number: A58833; MUID:96291862; PMID:8663314

A;Accession: A58833
 A;Molecule type: mRNA
 A;Residues: 1-352 <RAP>
 A;Cross-references: GB:U54994; NID:g1457945; PIDN:AAC50598.1; PID:g1457946
 C;Comment: This is a receptor for chemokines MIP-1alpha (see PIR:A30574), MIP-1beta (see PIR:A30574), and dual-tropic strains of HIV-1 bind to a complex of chemokine receptors and dual-tropic strains of HIV-1 bind to a complex of chemokine receptors.
 C;Genetics:
 A;Gene: GDB:CMKRS5; CCR5; CCR-5; CC-CR-5; CCR5; ChemR13
 A;Cross-references: GDB:I230510; OMIM:601373
 A;Map position: 3p21-3p21
 C;Function:
 A;Description: G protein-coupled receptor for chemokines MIP-1alpha, MIP-1beta and RANTES.
 A;Note: Probably acts to control granulocyte proliferation and differentiation.
 C;Keywords: AIDS; G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane protein; vertebate rhodopsin
 F;32-56/Domain: transmembrane #status predicted <TM1>
 F;67-87/Domain: transmembrane #status predicted <TM2>
 F;103-124/Domain: transmembrane #status predicted <TM3>
 F;142-166/Domain: transmembrane #status predicted <TM4>
 F;193-218/Domain: transmembrane #status predicted <TM5>
 F;236-257/Domain: transmembrane #status predicted <TM6>
 F;285-300/Domain: transmembrane #status predicted <TM7>
 F;20-269,101-178/Disulfide bonds: #status predicted
 F;268/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F;336,337,342/Binding site: phosphate (Ser) (covalent) #status predicted
 F;340,343/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 42.9%; Score 379; DB 2; Length 352;
 Best Local Similarity 35.0%; Pred. No. 8.1e-29;
 Matches 99; Conservative 7; Mismatches 27; Indels 150; Gaps 7;

QY 20 MDYQVSSPIVDINVTSEPCQKINVKQIAA----- 49
 DB 1 MDYQVSSPIVDINVTSEPCQKINVKQIAAARLLPPLYSLSVIFGFGNMLVILLINCR 60

QY 50 -----YKGLC----- 83
 DB 61 LKSMTDIYLLNLATISDLFELLTPFVAHAAAQWDFGNTMC--QLLTGLYFIFGFFSGIFF 118

QY 84 -----LCTRSQEGHLHYC 97
 DB 119 IILLTDIYLVAVHVAFLKARTVFGVTSVITVVAVFASLPGLIFTRSQEGHLHYC 178

QY 98 SSHFPYSQYQFQKNTQTLK-----HORVHG----- 124
 DB 179 SSHFPYSQYQFQKNTQTLKIVILGLVPLLVVICYSGLKTLRCRNEKRRHRAVRLIF 238

QY 125 -----GGSYKGLC-----QEFFGLNCCSSNRDLGHQV 154
 DB 239 TIMTVYFLFMAPYNIVLLNTFQEFFGLNCCSSNRDLQAMQV 281

RESULT 2
 157505
 zinc finger protein - mouse
 C;Species: Mus musculus (house mouse)
 C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 05-Nov-1999
 C;Accession: 157505
 R;Ennault-Lange, M.; Kress, M.; Hamer, D.
 Mol. Cell. Biol. 10, 418-421, 1990
 A;Title: A gene that encodes a protein consisting solely of zinc finger domains is preferentially expressed in the developing mouse brain.
 A;Reference number: 157505; MUID:90097859; PMID:2104662
 A;Accession: 157505
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-201 <RES>
 A;Cross-references: GB:M32057; NID:g199766; PIDN:AAA39729.1; PID:g199767
 C;Genetics:
 A;Gene: MOK2

Query Match 17.4%; Score 154; DB 2; Length 201;
 Best Local Similarity 37.0%; Pred. No. 1.7e-07;
 Matches 44; Conservative 13; Mismatches 44; Indels 18; Gaps 6;

QY 45 KQIAAYKCGLCAAQWDFGNTMCQHORVHGHHHSYKCGLCCTR--SQKEGLHYTCSSH- 101
 DB 60 KPACCECGMSFSQR-----SNLHQRVH-TGERPKYKCGCGKGFSSQSNLHHRCTHTG 114
 QY 102 --PYSQYQFQKNTF---QTLKHQRVHGCGGYSYKCGLCQEFFGLNCCSSNRDLGHQVH 155
 DB 115 EKPQYCECGKGFSSQSDLIHLRVHTGEKPYKCGCGGQF-----SQSKLLIHRVH 168

RESULT 3
 137570
 zinc finger protein - human (fragment)
 C;Species: Homo sapiens (man)
 C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 01-Dec-2000
 C;Accession: 137570
 R;Abrink, M.; Aveskog, M.; Hellman, L.
 DNA Cell Biol. 14, 125-136, 1995
 A;Title: Isolation of cDNA clones for 42 different Kruppel-related zinc finger proteins
 A;Reference number: 137566; MUID:95169271; PMID:7665130
 A;Accession: 137570
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-693 <RES>
 A;Cross-references: EMBL:X78927; NID:g498726; PIDN:CAA55527.1; PID:g498727
 C;Genetics:
 A;Gene: HZFA
 C;Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology

Query Match 16.3%; Score 144; DB 2; Length 693;
 Best Local Similarity 33.3%; Pred. No. 5.1e-06;
 Matches 39; Conservative 15; Mismatches 33; Indels 30; Gaps 7;

QY 40 QKINVKQIAAYKCGLCAAA-QWDFGNTMCQHORVHGHHHSYKCGLCCTRQSGEGLHYTCS 98
 DB 554 QKVTG-KPYKCGCGKGFMSLNDM--HQRVH-TGEKPYTCGAC-----GKHFSA 603

QY 99 SHFPYSQYQFQKNTQTLKHQRVHGCGGYSYKCGLCQEFFGLNCCSSNRDLGHQVH 155
 DB 604 S-----SLQLHQSVHTGEKPYKCDVCGKVF-----SRSSQLQYHRRVH 641

RESULT 4
 T46469
 hypothetical protein DKFZp434G1930.1 - human (fragment)
 C;Species: Homo sapiens (man)
 C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
 C;Accession: T46469
 R;Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, January 2000
 A;Reference number: 223034
 A;Accession: T46469
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-319 <AAA>
 A;Cross-references: EMBL:AL137483
 A;Experimental source: adult testis; clone DKFZp434G1930
 C;Genetics:
 A;Note: DKFZp434G1930.1

Query Match 16.2%; Score 143.5; DB 2; Length 319;
 Best Local Similarity 32.8%; Pred. No. 2.7e-06;
 Matches 38; Conservative 15; Mismatches 44; Indels 19; Gaps 7;

QY 50 YKGLCAAAQWDFGNTMCQHORVHGHHHSYKCGLCCTRQSGEGLHYTCSHF-----P 102
 DB 9 YKNECEKT-FSHRSSLLSHQRIH-TGEKPYKNECEKAFSNS--STLIKHLRVHTGKXP 64

QY 103 YSQYQFQKNTF---QTLKHQRVHGCGGYSYKCGLCQEFFGLNCCSSNRDLGHQVH 155
 DB 65 YRCRCGKAFSQCSLTAVHORIHTEKLYKCGCEKAF---NCRA--KLHRHQRIH 115

RESULT 5

I53869
zinc finger protein - mouse
C:Species: Mus musculus (house mouse)
C>Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 01-Dec-2000
C:Accession: I53869
R:Brady, J.P.; Platiogorsky, J.
Gene 149, 299-304, 1994
A:Title: A mouse cDNA encoding a protein with zinc-fingers and a KRAB domain shows similarity to the human zinc finger protein ZNF134
A:Reference number: I53869; MUID:95047492; PMID:7959006
A:Accession: I53869
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-555 <RES>
A:Cross-references: GB:128167; NID:9758660; PIDN:AAA67545.1; PID:9758661
C:Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology

Query Match 16.0%; Score 141; DB 2; Length 555;
Best Local Similarity 23.8%; Pred. No. 8e-06;
Matches 50; Conservative 25; Mismatches 47; Indels 88; Gaps 11;

QY	23	QYSSPIYDINYTSBPCQ-----KINVKQIAAAYKGLCAAQWDFGNTMCOH	69
DB	303	RVCTP---VAPYCEQCEKSLLDVQHLMSHVKVHTRE-RPYNCETCGSA-FSQASHLQDH	357
QY	70	QRVH-----GHHHS-----YKGLCTRSQKEGLHYTCSSHF-	101
DB	358	QLHTGCKPKKCDACGKFSRSSHLRSHQVHTGEKPYKCGCGKS-----FICSSNLY	411
QY	102	-----PSYQYQFWKNF-----QTLKIHQHVHGGGSKGLCQEPFGLNN-----	141
DB	412	IHORVHTGEKPYKCDVCGKEFRSPSLQAHQGHGTGKSYVCTMGKGYTLNSLQVHLR	471
QY	142	-----C-----SSSNRLDGHQVH	155
DB	472	VHTGKPYSCDVCGKGFSSRSQLOSHQVH	501

RESULT 6

F31201
GLI-related finger protein HKR4 - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 17-Mar-1994
C:Accession: F31201
R:Ruppert, J.M.; Kinzler, K.W.; Wong, A.J.; Bigner, S.H.; Kao, F.T.; Law, M.L.; Seunanez, M.L.; Cell. Biol. 8, 3104-3113, 1988
A:Title: The GLI-Kruppel family of human genes.
A:Reference number: A93103; MUID:89096896; PMID:2850480
A:Accession: F31201
A:Molecule type: DNA
A:Residues: 1-223 <RUP>
C:Keywords: DNA binding; zinc finger

Query Match 15.7%; Score 138.5; DB 2; Length 223;
Best Local Similarity 30.7%; Pred. No. 5.6e-06;
Matches 42; Conservative 13; Mismatches 39; Indels 43; Gaps 7;

QY	50	YKGLCAAQWDFGNTMCOHQHVHGHHSYKGLCTRSQKEGLHYTCSSHPYQYQFW	109
DB	96	YEQCCGGA-FSSHSHFTQHLRIH-NGEKPYKCGECVRHQH-LH---TGERPYACSG	148
QY	110	KNF---QTLKIHQHVHGGGSKY-----CGLCQEPFG	138
DB	149	KAFIWSVLTEHQRIHTGKPYECDCGKAFRGRSHFRHLRHTHTGKPFACGACGKAFG	208
QY	139	LNNCSSNRLDGHQVH	155
DB	209	-----QSSQLIQHVH	220

RESULT 7

I38599
zinc finger protein ZNF134 - human

C:Species: Homo sapiens (man)
C>Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 05-Nov-1999
C:Accession: I38599
R:Tommerup, N.; Vissing, H.
Genomics 27, 259-264, 1995
A:Title: Isolation and fine mapping of 16 novel human zinc finger-encoding cDNAs identical to the human zinc finger protein ZNF134
A:Reference number: A57785; MUID:96044430; PMID:7557990
A:Accession: I38599
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-348 <RES>
A:Cross-references: EMBL:U09412; NID:9488552; PIDN:AAC50253.1; PID:9488553
C:Genetics:
A:Gene: ZNF134
A:Cross-references: GDB:137033
A:Map position: 19q13.4-19q13.4

Query Match 15.4%; Score 136.5; DB 2; Length 348;
Best Local Similarity 27.3%; Pred. No. 1.4e-05;
Matches 45; Conservative 13; Mismatches 46; Indels 61; Gaps 9;

QY	50	YKGLCAAQWDFGNTMCOHQHVHGHHSYKGLC-----TRSQEGLH-----YT	96
DB	97	YKCEGGA-FSRKDTLVQHQRIHS-GEKPYECGKAFSRKATLVQHQRIHTGERPYE	154
QY	97	CS-----SHFPYQYQFWKNF---QTLKIHQHVHGGGSKGLC	133
DB	155	CSECGKTPSRKDNLTQHKRIHTGEMPKYKCEGKYFSHSHNLVHQVHNGARPYKCSDC	214
QY	134	QEPF-----GLN---NCSSSNRLDQ-----HORVH	155
DB	215	GKVFRIKSLVQHSIHTGENPNYDCDGKSGHGYTLIKHQRIH	259

RESULT 8

T12527
hypothetical protein DKFP434N043.1 - human (fragment)

C:Species: Homo sapiens (man)
C>Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 01-Dec-2000
C:Accession: T12527
R:Mambutti, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, June 1999
A:Reference number: Z17524
A:Accession: T12527
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-386 <WAM>
A:Cross-references: EMBL:AL080143
A:Experimental source: adult testis; clone DKFP434N043
C:Genetics:
A:Note: DKFP434N043.1
C:Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology

Query Match 15.4%; Score 136.5; DB 2; Length 386;
Best Local Similarity 20.2%; Pred. No. 1.5e-05;
Matches 47; Conservative 22; Mismatches 73; Indels 91; Gaps 7;

QY	6	AALSLMLVTALGSDMDYQVSSPIYDINYTSBPC-----QKINVKQ	46
DB	43	STLKCHSVHAREKQGFVSGKILDQNPQEKCKPKCNKCEKTFSCSKYLTQHERIHRG	102
QY	47	IAAYKGLCAAQWDFGNT---MCOHQVHGH-----HHH	78
DB	103	VKQEPEDCCGKA---FGQSTRLIHQHRSRVLRYKMGQGGKAISSASLIKQSPHTEH	159
QY	79	YKGLCTRSQKEGLHYT-----CSSHPYQYQ-----	106
DB	160	PFKNECGKTFSSHASHLSKQLIHAGENPFKSKCDRVTQRYLVQHERHARKKPLVC	219
QY	107	----QFWKHFOTLKIHQHVHGGGSKGLCQEPFGFLANNCSSNRLDGHQVH	155
DB	220	NECGKTFROSSCLSKHQRIHSGEKPYVCDYCGKAFGL-----SALVHRHQRIH	267

Best Local Similarity 100.0%; Pred. No. 9.3e+05; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
| | | | |

Db 1 GGGGS 5

RESULT 84

AAU76044
ID AAU76044 standard; Peptide; 6 AA.

XX
AC AAU76044;

DT 08-MAY-2002 (first entry)

Synthetic linker moiety peptide sequence #2.

XX Human; A Disintegrin And Metalloproteinase with Homology to ADAM9;
DE ADAM-H9; cytosolic; antiinflammatory; osteopathic; vasotropic;
XX thrombolytic; vulnaray; antiallergic; antibacterial; virocid;
KW procoagide; analgesic; angiogenesis; endothelial cell migration;
KW cellular disorder; restenosis; thrombosis; cardiovascular disorder;
XX chronic pain condition; haematological disorder; cancer; tumour.

XX Synthetic.

XX Key Location/Qualifiers

PH Misc-difference 6 /label= OTHER

FT /note= "OTHER= Xaa is one or more repeats of GGGGS"

FT WO200210405-A2.

XX 07-FEB-2002.

XX 27-JUL-2001; 2001WO-US23709.

XX 28-JUL-2000; 2000US-221838P.

PR 09-APR-2001; 2001US-282550P.

XX (IMMV) IMMUNEX CORP.

XX Black RA, Poindexter K, Mosley BA, Dubose RF, Wiley SR;

XX WPI; 2002-188750/24.

XX Novel substantially purified human polypeptide having homology to A
PT Disintegrin and Metalloproteinase domain (ADAM) polypeptide family,
PT termed ADAM-H9, for treating ocular and inflammatory diseases,
PT osteoporosis, cancer -

XX Disclosure; Page 13; 72pp; English.

XX The present invention relates to a new human polypeptide having homology
XX to A Disintegrin And Metalloproteinase domain (ADAM) polypeptide family,
XX termed ADAM-H9. The polypeptide of the invention is useful for
XX identifying an agent that modulates an activity of the polypeptide, in a
XX inhibiting angiogenesis in a mammal, for modulating angiogenesis in a
XX tissue, for modulating endothelial cell migration in vitro or in vivo,
XX for inhibiting the binding of an integrin to a ligand and for modulating
XX the binding of an integrin to a ligand in a mammal afflicted with a
XX condition such as ocular disorders, malignant and metastatic conditions,
XX inflammatory diseases, osteoporosis, accelerated bone resorption
XX disorders, restenosis, inappropriate platelet activation, recruitment or
XX aggregation, thrombosis and a condition requiring tissue repair or wound
XX healing. The polypeptide used for treating the above conditions is a
XX multimer (a dimer or trimer) which comprises an FC polypeptide, a leucine
XX zipper or a peptide linker. The invention is useful for treating
XX disintegrin-associated disorders and conditions and is also useful for
XX treating medical conditions and diseases associated with cell-cell and
XX cell matrix interactions, endothelial migration, angiogenesis,
XX inflammation, allergy, reproductive, neurological and vascular

CC conditions, bacterial, viral or protozoal infections, cardiovascular
CC disorders such as myocardial infarction, chronic pain conditions,
CC endocrine system disorders, gastrointestinal system disorders,
CC genitourinary system disorders, anaemia, haematological disorders and
CC oncologic conditions such as cancer. The present amino acid sequence
CC represents one of several (AAU76043-AAU76049) linker moiety peptides of
CC the invention. These linkers were used in the invention to join an
CC ADAM-H9 disintegrin domain to an FC domain or leucine zipper domain.

XX Sequence 6 AA;

Query Match 100.0%; Score 5; DB 23; Length 6;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5

Db 1 GGGGS 5

RESULT 85

AAU82046

ID AAU82046 standard; peptide; 6 AA.

XX AC AAU82046;

XX 09-APR-2002 (first entry)

XX T-cell specific binding ligand peptide construct peptide spacer #1.
DE T-cell specific binding ligand peptide construct peptide spacer #1.
XX

XX T-cell binding ligand; TCBL; peptide G'; human MHC class II beta chain;
KW peptide J; human beta-2-microglobulin; HIV-1; TCBL peptide construct;
KW immunological disorder; immune response; human immunodeficiency virus;
KW herpes simplex virus infection; HSV; malaria; tuberculosis; cancer; CEA;
KW acquired immunodeficiency syndrome; AIDS; allergy; autoimmune disease;
KW autoimmune myocarditis; cytostatic; antiinflammatory.

XX Synthetic.

XX WO200189286-A2.

XX 29-NOV-2001.

XX 24-MAY-2001; 2001WO-US16793.

XX 24-MAY-2000; 2000US-206548P.

XX (CELS-) CEL-SCI CORP.

XX Zimmerman DS, Sarin PS;

XX WPI; 2002-083037/11.

XX New T cell binding ligand peptide for treating immunological disorders
XX such as herpes simplex virus, tuberculosis, cancers, acquired
XX immunodeficiency syndrome and allergies -

XX Disclosure; Page 18; 110pp; English.

XX The present invention relates to novel T-cell binding ligand (TCBL)
XX peptides (e.g. peptide G' (modified human MHC class II beta chain
XX peptide G, peptide J (human beta-2-microglobulin peptide) and HIV-1
XX peptides) and TCBL peptide constructs for treating immunological
XX disorders). The peptide constructs are useful for eliciting a cellular
XX immune response in a human patient. The method comprises administering
XX the peptide construct to the patient preferably in combination with an
XX immune response adjuvant. The peptide constructs in the form of
XX conjugated peptides are useful for eliciting a cellular immune response
XX in a patient exposed to or at risk for exposure to the human
XX immunodeficiency virus (HIV). The TCBL peptides are useful for treating a
XX patient suffering from an immunological disorder such as herpes simplex
XX virus (HSV) infection, malaria, tuberculosis, cancers, acquired
XX immunodeficiency syndrome (AIDS), allergies, autoimmune diseases

CC their modulators are also useful for treating pancreatic disorders,
CC heart disorders, pituitary-related disorders, disorders of the adrenal
CC cortex, thyroid, placenta, gastric disorders, disorders of the adrenal
CC disorders of the skeletal muscle, hepatic disorders, pulmonary disorders,
CC hepatitis), renal disorders (e.g. glomerulonephritis), testicular
CC disorders, uterine disorders, disorders of the brain, central nervous
CC system disorders including Alzheimer's disease, Parkinson's disease,
CC multiple sclerosis, brain cancers, eye disorders (e.g. glaucoma,
CC cataract), prostate and ovarian disorders, intestinal disorders,
CC colonic, and leukocytic disorders. ABG78024-ABG78125 represent ITALY,
CC Lor-2, STRIFE, TRASH, BDSF, LRSG and STMST amino acid sequences and
CC related peptides of the invention.

SQ Sequence 6 AA;
Query Match 100.0%; Score 5; DB 23; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|
|
|
|
|
Db 1 GGGGS 5

RESULT 82
ABB84094
ID ABB84094 standard; peptide; 6 AA.

AC ABB84094;

XX 09-SEP-2002 (first entry)

DT Human scdb CEAGAL PA & PB linker peptide (wild type).

DE Antibody; multimer; cytostatic; detection; immunoassay; tumour.

KW Homo sapiens.

OS DE10060140-A1.

PN 06-JUN-2002.

PD 04-DEC-2000; 2000DE-1060140.

XX 04-DEC-2000; 2000DE-1060140.

PR (VECT-) VECTRON THERAPEUTICS INT AG.

PA Kontermann R;

PI WPI; 2002-520984/56.

DR New multimer of single-chain antibodies, useful e.g. for diagnosis or
XX drug delivery, has four variable chains linked through peptides of
XX controlled length -

PS Example 1; Page 10; 16pp; German.

XX This invention describes a novel multimer containing at least two
XX antibodies, each, independently, comprising a polypeptide chain of
XX structure V1-PA-V2-PM-V3-PB-V4 where V1-V4 = variable domains; PA, PB
XX and PM = peptide linkers with PA and PB containing 0 or 1 amino acid. The
XX products of the invention have cytostatic activity. The multimers
XX described can be used for diagnostic detection, in essentially standard
XX immunoassay methods, and for binding at least one component and/or fused
XX protein or peptide to a cell, especially to deliver a toxin or for
XX infection, transformation or transfection of the cell. (I), or nucleic
XX acid (II) encoding it, or vectors containing (II), are useful for
XX treating tumours. The multimers are relatively small (about 110 kDa for a
XX dimer) but have multiple binding sites to ensure high binding affinity.
XX This sequence represents a wild-type PA and PB type linker peptide
XX described in the method of the invention.

SQ Sequence 6 AA;
Query Match 100.0%; Score 5; DB 23; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|
|
|
|
|
Db 2 GGGGS 6

RESULT 93
AAO14388
ID AAO14388 standard; Peptide; 6 AA.

AC AAO14388;

XX 24-MAY-2002 (first entry)

DT Linker peptide for use in a fusion protein #2.

DE Human; metalloproteinase-like/disintegrin-like protein; MPD;
XX angiogenesis; endothelial cell migration; integrin ligand binding;
XX ocular disorder; cancer; inflammation; osteoporosis; restenosis;
XX cytostatic; antiinflammatory; osteopathic; vasotropic; thrombolytic;
XX vulnary; antiallergic; antibacterial; virucide; protozoacide; cardiant;
XX antianaemic; analgesic; thrombosis; tissue repair; neural disorder;
XX infection; linker.

OS Synthetic.

FH Key Location/Qualifiers

FT Misc-difference 6 /label= Xaa

FT /note= "Xaa-one or more repeats of the present sequence"

XX WO200210406-A2.

XX 07-FEB-2002.

XX 27-JUL-2001; 2001WO-US23734.

XX 28-JUL-2000; 2000US-221838P.

XX (IMMV) IMMUNEX CORP.

XX Dubose RP, Wiley SR, Black RA;

XX WPI; 2002-195965/25.

XX Novel purified human metalloproteinase-disintegrin polypeptide, useful
XX for treating ocular and inflammatory disorders, osteoporosis, cancer,
XX restenosis, thrombosis, and chronic pain conditions -

PS Disclosure; Page 20; 101pp; English.

XX The present invention provides a number of human metalloproteinase-like
XX /disintegrin-like (MPD) proteins and their fragments. These can be used
XX to identify treatments for ocular disorders, malignant and metastatic
XX conditions, inflammatory diseases, osteoporosis and other conditions
XX mediated by accelerated bone resorption, restenosis, inappropriate
XX platelet activation, recruitment or aggregation, thrombosis, conditions
XX requiring tissue repair or wound healing, endothelial migration,
XX angiogenesis, allergies, reproductive, neurological and vascular
XX conditions, infections, cardiovascular disorders such as myocardial
XX infarction, chronic pain conditions, endocrine system disorders,
XX gastrointestinal system disorders, genitourinary system disorders,
XX anaemia and haematological disorders. The present sequence is a linker
XX peptide described in the exemplification of the invention.

SQ Sequence 6 AA;

Query Match 100.0%; Score 5; DB 23; Length 6;